DESIGNING OPTIMAL DEMAND-RESPONSIVE TRANSPORTATION FEEDER SYSTEMS AND COMPARING PERFORMANCE IN HETEROGENEOUS ENVIRONMENTS

A Thesis
Presented to
The Academic Faculty

by

Derek Lee Edwards

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Electrical and Computer Engineering

Georgia Institute of Technology
August, 2014

DESIGNING OPTIMAL DEMAND-RESPONSIVE TRANSPORTATION FEEDER SYSTEMS AND COMPARING PERFORMANCE IN HETEROGENEOUS ENVIRONMENTS

Approved by:

Dr. George Vachtsevanos, Advisor
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Steve Dickerson
George W. Woodruff School of Mechanical Engineering
Georgia Institute of Technology

Dr. Kari Watkins
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. David Taylor
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Jennifer Michaels
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Chuanyi Ji
School of Electrical and Computer Engineering
Georgia Institute of Technology

Date Approved: April 2014

To Ife.

ACKNOWLEDGEMENTS

My time at Georgia Tech allowed me to work with an amazing set of colleagues. All of which were instrumental in helping me to complete this thesis. First, I'd like to thank Kari Watkins and the team at the Urban Transportation Information Lab (UTIL) for their hands-on support helping a computer engineer find his way in the world of transportation. The time I spent at UTIL were the most intellectually fulfilling of my time at Georgia Tech . Without the wealth of domain knowledge and hands-on guidance that I had access to at UTIL, this research may have never advanced beyond an interesting hobby.

I'd like to thank Steve Dickerson, Aarjav Trivedi, and Arun Kumar-Elangovan at RideCell for giving me my first hands-on experience in transportation. I learned more practical programming skills during my time at RideCell than any other period while at Georgia Tech. Without the experiences gained at RideCell, I would not have the technical skills needed to perform the work required for my research.

I'd like to thank David Taylor and Ryan Melsert for their efforts in leading the Georgia Tech EcoCAR team. Our marathon bi-weekly team meetings with Dr. Taylor gave me a unique combination of academic rigor combined with practical application.

Finally, I'd like to thank Dr. Vachtsevanos and my friends from the Intelligent Control Systems Lab at Georgia Tech for helping me get started in the scary world of academic research. Without the initial support from Dr. V and encouragement from my colleagues, Doug Brown, Brian Bole, and countless others, I would never have made it out of the starting blocks in my academic pursuits.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF ABBREVIATIONS

BRT .... Bus Rapid Transit

DARP .... Dial-a-Ride Problem

DRT .... Demand-Responsive Transportation

FRT .... Fixed-Route Transportation

GTFS .... General Transit Feed Specification

IDARP .... Integrated Dial-a-Ride Problem

MARTA .... Metro Atlanta Transportation Authority

NITS .... Network-Inspired Transportation System

OSRM .... Open Source Routing Machine

OTP .... OpenTripPlanner

TSP .... Traveling Salesman Problem

VMT .... Vehicle Miles Traveled

SUMMARY

The goal of this research is to develop a method of objectively comparing and optimizing the performance of demand-responsive transportation systems in heterogeneous environments. Demand-responsive transportation refers to modes of transportation that do not follow fixed routes or schedules, including taxis, paratransit, deviated-route services, ride sharing as well as other modes. Heterogeneous environments are transportation environments in which streets do not follow regular patterns, passenger behavior is difficult to model, and transit schedules and layouts are non-uniform. An example of a typical heterogeneous environment is a modern suburb with non-linear streets, low pedestrian activity, and infrequent or sparse transit service. The motivation for this research is to determine if demand-responsive transportation can be used to improve customer satisfaction and reduce operating costs in suburban and low-density urban areas where fixed-route transportation may be inefficient.

This research extends existing comparison and optimization techniques that are designed to work in homogeneous environments. Homogeneous environments refer to transportation systems where the streets follow regular and repeating patterns, passengers are evenly distributed throughout the system, and the transit system is easily modeled. The performance of systems with these characteristics can be approximated with closed-form analytical expressions representing passenger travel times, vehicle distances traveled, and other performance indicators. However, in the low-density urban areas studied in this research, the street patterns and transit schedules are irregular and passenger behavior is difficult to model. In these areas, analytical solutions cannot be found. Instead, this research develops a simulation-based approach to compare and optimize performance in these heterogeneous environments. Using widely-

available route-planning tools, open-source transit schedules, and detailed passenger data, it is possible to simulate the behavior of transit vehicles and passengers to such an exacting degree that analytical solutions are not needed.

A major technical contribution of this research is the development of a demand-responsive transportation simulator to analyze performance of demand-responsive systems in heterogeneous environments. The simulator combines several open-source tools for route planning with a custom-built demand-responsive vehicle and passenger-itinerary optimizer to simulate individual vehicles and passengers within a large system. With knowledge of the street network, the transit schedule, passenger locations, and trip request times, the simulator will output exact passenger transit times, passenger travel distances, vehicle travel distance, and other performance indicators for a particular transportation setup in a given area.

The simulator is used to develop a method of comparing various demand-responsive and fixed-route systems. By predefining a set of performance indicators, such as passenger travel time and operating cost, the simulator can be used to ascertain the performance of a wide array of transportation systems. Comparing the weighted cost of each type of system permits a transportation engineer or planner to determine what type of system will provide the best results in a given area.

The simulator is extended to assist in optimization of the demand-responsive transportation system layout. A key problem that needs to be solved when implementing a demand-responsive system is to determine the size, shape, and location of the demand-responsive coverage areas, i.e., the areas in which passengers are eligible for demand-responsive transportation. Using a particle swarm optimization algorithm and the simulation-based comparison technique, the optimal size and shape for a demand-responsive coverage area can be determined.

The efficacy of the comparison and optimization techniques is demonstrated within the city of Atlanta, GA. It is shown that for certain areas of the city of Atlanta, demand-responsive transportation is more efficient than the currently implemented fixed-route system. Depending on the objective of the transportation planner, passenger satisfaction as well as operating costs can be improved by implementing a demand-responsive system in certain low-density areas.

The techniques introduced in this research, and the simulation tool developed to implement those techniques, provide a repeatable, accurate, and objective method with which to optimize and compare demand-responsive transportation systems in heterogeneous environments.

CHAPTER 1

INTRODUCTION

The objective of this research is to develop a method of optimizing the design of demand-responsive transportation feeder systems and to develop a method of objectively comparing demand-responsive transportation service and fixed-route service in heterogeneous environments. The motivation for this research is to determine if demand-responsive transportation or a combination of demand-responsive and fixed-route transportation can be used to improve customer satisfaction and reduce operator costs in low-density urban areas where fixed-route transit may not be efficient. In this research, demand-responsive transportation refers to point-to-point transportation systems in which a passenger makes a trip request and a vehicle is dispatched to handle that trip. The vehicle may be a shared-ride vehicle, such as a dial-a-ride service, or the vehicle may be a single-ride vehicle, such as a taxi. Fixed-route transportation refers to traditional bus and rail service with inflexible routes and schedules.

This research extends the current state-of-the-art of designing and comparing demand-responsive systems. The current methods are designed for homogenous environments. Homogeneous environments refer to study areas where passenger data, street layouts, and transit schedules are easily represented by closed-form solutions. In homogeneous systems, passenger trip-request rates and passenger spatial distribution can be accurately modeled by probability distribution functions, street networks follow regular grid or ring-radial patterns, and transit service is often assumed to follow a regular and unchanging schedule. However, the motivation for this research is to study the effects of demand-responsive transportation in low-density and suburban areas that lack many of the homogeneous characteristics needed to apply the existing optimization and comparison techniques.

This research seeks to develop a generic, simulation-based approach to study demand-responsive transportation in a variety of heterogeneous environments. Heterogeneous environments refer to road systems that do not necessarily follow a particular pattern, non-uniform distributions of passengers in both time and space, and complex transit schedules that are difficult to codify into a closed-form representation. In previous work, certain assumptions are made in order to simplify the problem to the point that closed-form representations can model the demand-responsive and fixed-route system behavior. In this research, no assumptions are made about passenger arrival rates, passenger locations, street layouts, or transit schedules.

The complex nature of heterogeneous environments makes them difficult to model with closed-form expressions. Therefore, this research takes a data-driven approach to optimization and leverages recent efforts to open transportation data and route-optimization software to researchers. The mass adoption of a uniform transit schedule format known as general transit feed specification (GTFS), as well as availability of open source mapping data and routing tools, allows for complex simulations to be run that will model the precise behavior of passengers and vehicles within demand-responsive and fixed-route systems. The availability of these tools and data allow for any environment to be studied without the need to simplify the problem.

This research combines available open-source routing tools, passenger survey data, street map data, and transit schedule data to build a custom demand-responsive transportation simulator. The simulator is used to analyze demand-responsive performance within heterogeneous environments and assists with designing optimal demand-responsive feeder zones. The simulator, which is thoroughly examined in Chapter 4, is highly modular and is intended to allow transit operators to study the effects of demand-responsive transportation in a variety of settings. The simulator provides the backbone for the comparison and optimization algorithms introduced by this research.

This dissertation is organized as follows. Chapter 1 discusses the motivation of this research by demonstrating the feasibility of using demand-responsive transportation in low-density areas where fixed-route transportation may be inefficient. Chapter 2 provides a thorough literature review of the existing demand-responsive technologies and point out areas of improvement that this research will address. Chapter 3 introduces the network-inspired transportation system framework. Chapter 4 discusses the operation of the demand-responsive simulator built for this research. Chapter 5 demonstrates a method of objectively comparing fixed-route transportation and demand-responsive transportation in heterogeneous settings. Chapter 6 introduces a particle swarm optimization technique to optimize the shape and size of demand-responsive feeder zones in heterogeneous environments. Finally, the importance of this research, conclusions, and future work are summarized in Chapter 7.

1.1 Motivation

This research develops a method of optimally designing demand-responsive feeder systems and introduces a way to objectively compare demand-responsive and fixed-route feeder systems in heterogeneous environments. However, in order to understand why such methods are needed, it is important to understand how demand-responsive transportation can be used to improve the performance of transit systems in low-density areas.

A major problem in increasing ridership on public transit has been reaching destinations outside of dense urban areas. In medium and low density areas, it may not be economically viable for transit to run frequent service within easy walking distance to every location [1]. As a result, ridership in these areas remains low as compared to other forms of transportation, especially personal automobile transportation [2].

There is a strong relationship between population density and ridership [1]. Higher population density often correlates with higher transit ridership. Cities and areas with low density of riders are challenged to provide high quality of service transportation. Consider two cities with very different development patterns and transit usage characteristics: Atlanta and New York. In Atlanta, GA, which has a population density that is 85% lower than that of New York [3], a bus stop or rail station will be within walking distance to only 15% as many people as a bus stop or rail station in New York. Due to this, a transit system in Atlanta would need to operate over a geographic area that is 6.7 times larger than a transit system in New York in order to reach the same number of potential passengers. Balancing the costs associated with providing fixed-route transportation is often a tradeoff between covering a large area and providing frequent and fast service [4]. Therefore, in order to provide resources, in terms of drivers and vehicles, over a large area, often average headways will be increased to save costs. This is a problem faced by the Metropolitan Atlanta Rapid Transit Authority (MARTA). In order to reach stops across a two-county footprint with a relatively low-population density of 2,046 people per square mile, average headways are kept high in order to keep costs low. Between 7:00 AM and 7:00 PM, the average headway across the MARTA service area is 25 minutes. According to the Transit Capacity and Quality of Service Manual issued by the Transportation Research Board, this is an average level of service grade of D. Level of service D is considered to be "unattractive to choice riders" [5].

While much of the MARTA coverage area has a relatively low ridership and density, there are corridors within the city that have relatively high ridership and frequent service. Figure 1-1 shows a heat map of the passenger origins and destinations within the MARTA coverage area.

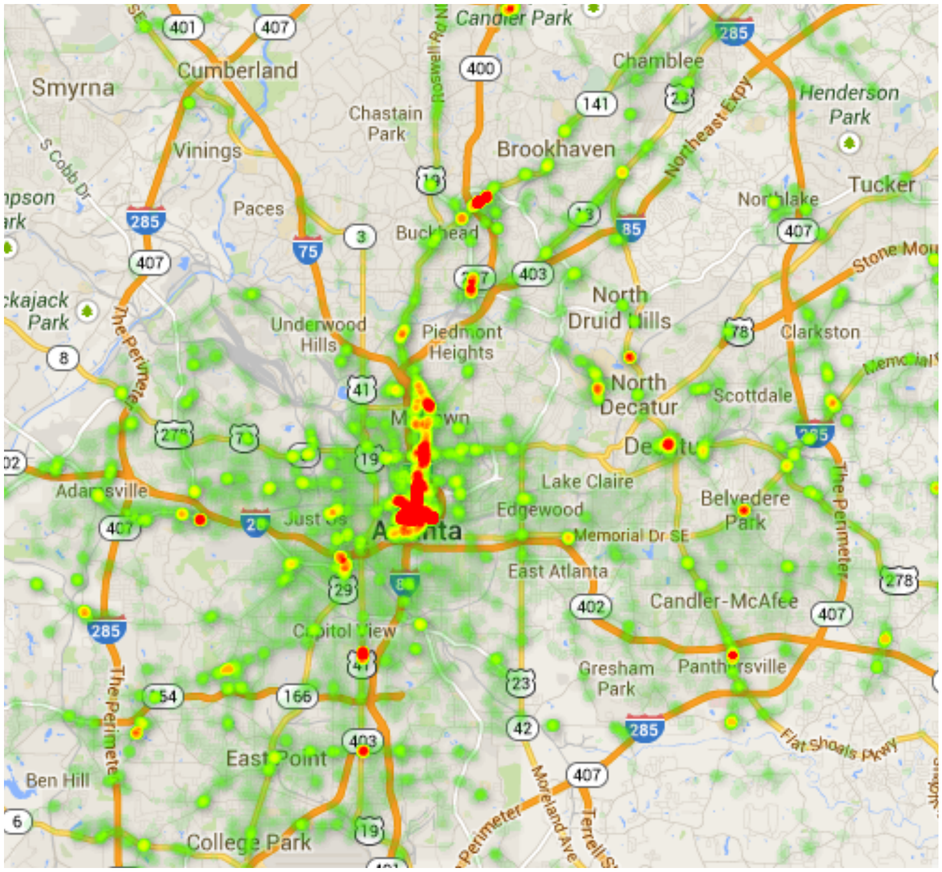

Figure 1-1: Heat map of MARTA passenger origins and destinations.

The areas in the center of the map, Midtown and Downtown Atlanta, have the highest concentrations of passenger origins and destinations. These areas correlate highly with the bus and rail stops with the most frequent service, bus route 110 and the rail routes operate at 15 minute headways, which is on the high end of service level C [5]. Service level C is defined as fixed-route service operating headways of 15-20 minutes.

It is hypothesized that the high frequency rail and bus lines located in the areas of the city with high ridership can be combined with a demand-responsive service operating in the areas of the city with lower ridership and higher headways. The purpose of the demand-responsive service would be to provide first-mile and last-mile connectivity between rail stations and areas of low ridership. Often, the most time-consuming portion of a passenger's trip is waiting on low-frequency buses in low-density areas of the city. If some of these low-frequency bus routes were replaced with a demand-responsive service, total door-to-door travel times could be reduced. This research will test this hypothesis for various locations and times within the city of Atlanta.

Historically, demand-responsive transportation has been limited to rural routes, where ridership is below the minimum demand needed to justify fixed-route service, or demand-responsive services is offered as an alternative to passengers who are unable to use the fixed-route buses due to mobility issues. The logistical and communication challenges encountered when requesting trips and arranging optimized routes have led to many demand-responsive services to require advance notice in order to make trip requests. Often, this advance noticed must be placed more than one day in advance [6]. However, several technological innovations in recent years have led to newfound interest in demand-responsive service, specifically inexpensive GPS devices on vehicles and pervasive mobile phone technology [7]. Instead of arranging rides hours or days ahead of time, mobile communication devices such as tablets and GPS units allow easy communication between passengers, drivers, and dispatchers. Because of this, trips can now be requested instantly via mobile phone or internet and drivers can receive new itineraries instantly each time a trip is assigned to their vehicle. Reducing the logistical hurdles to arranging routes and requesting rides has made demand-responsive transportation a viable option to fixed-route for some transit operators.

Among the new adopters of demand-responsive transportation systems are universities, such as Georgia Tech [8], Duke University [9], and Case Western Reserve University [10] among others. These universities use demand-responsive transportation as a substitute for fixed-route service when demand is low, typically during the evening hours. In addition to universities, cities have implemented demand-responsive systems on a large scale for general public use. Cities such as Denver's Call-n-Ride [11], Helsinki, Finland,[12], Charleston, SC [13], and Glasgow, UK [14] among others have implemented demand-responsive services both as first and last-mile services as well as door-to-door services in low-demand areas.

This research seeks to provide tools to transportation engineers who are considering implementing demand-responsive transportation. Two difficult problems that transportation engineers encounter when considering demand-responsive options for their jurisdictions are determining which areas of the city are best suited for demand-responsive transportation and designing the demand-responsive feeder system to meet customer demands while minimizing costs to the transit operator. Current methods of comparison and feeder-system optimization, which are discussed in greater detail in Chapter 2, are designed for use in cities with gridded street systems and predictable passenger behavior. While these methods provide a great framework for comparing and optimizing demand-responsive transportation, the methods are not easily applied to many of the medium and low-density cities that would benefit from demand-responsive transportation.

CHAPTER 2

LITERATURE REVIEW

The two major objectives of this research are to develop a method of objectively comparing demand-responsive transportation services and fixed-route services in heterogeneous environments as well as to develop a method of optimizing the design of demand-responsive transportation feeder systems. In order to better understand how this research addresses current problems within the study of demand-responsive transportation, this chapter provides a review of the relevant research. The review includes a broad overview of demand-responsive transportation research as well as a focused investigation into current methods of comparing demand-responsive and fixed-route transportation as well as current methods of optimizing demand-responsive feeder systems. The shortcomings of current comparison and optimization techniques will be listed, motivating the development of the comparison and optimization techniques introduced in Chapters 4 and 5.

2.1 Review of Demand-Responsive Transportation

Demand-responsive transportation (DRT) refers to transportation that does not follow a fixed schedule or route. Examples of demand-responsive transportation include taxi services, ad hoc carpooling or slugging, jitney services, dial-a-ride services, and deviated-route transit. In order to focus the scope of this research, the work presented in this thesis focuses primarily on dial-a-ride services. However, future work is suggested to include additional forms of demand-responsive transportation.

In a dial-a-ride service, multiple passengers make trips requests and one or more vehicles are dispatched to meet those requests by a central dispatching agent. Multiple trips will often be handled by the same vehicle simultaneously in order to reduce costs to the operator. A common use of dial-a-ride services is to provide transit options to elderly or disabled persons as well as to provide service to rural areas that lack a fixed-route system [15] [16]. This research utilizes dial-a-ride as a means of providing transportation options to the general public in low-density urban areas without enough ridership demand to support frequent fixed-route transit.

The fundamental problem of any demand-responsive transportation system is that of assigning passengers to the optimal vehicle and selecting the optimal route for that vehicle in order to meet passenger demand while minimizing operator costs. In mathematics and optimization theory, this problem is known as the dial-a-ride problem (DARP). Specifically, the dial-a-ride problem consists of creating *m* optimal bus and van routes to service a set of *n* passengers, curb-to-curb, with *a priori* information of the passengers' origins and destinations. A thorough mathematical model of the DARP is presented by Cordeau and Laport in [15].

The DARP can be divided into several sub-categories relating to the number of buses used in the system, how and when passengers are added to the system, and the level of interaction between fixed and dynamic vehicle routes. A review and discussion of the advantages and disadvantages of these characterizations are given below.

2.1.1 Single Bus versus Multiple Bus Dial-a-Ride

The first characterization separates the dial-a-ride problem into single-bus systems and multi-bus systems. A single-bus dial-a-ride system is defined as demand-responsive transportation system in which all passengers are serviced between their origins and destinations by a single bus. The single-bus formation of the DARP is a similar to the classic traveling salesman problem (TSP). In the traveling salesman problem, a salesman must select the optimal route with which to visit *n* cities. The single-bus DARP is a variation of the TSP where the bus represents the salesman and the passenger locations and destinations represent the cities. The only additional constraint in the single-bus DARP is that a passenger's destination cannot be visited before that passenger's origin. In-depth studies of the single vehicle DARP is provided in [17] [18], and [19].

A more complex version of the DARP is the multi-bus DARP. In this version of the DARP, passengers requesting rides can be split among multiple buses. This additional degree of freedom present in the multi-bus DARP requires two optimization decisions to be made; assigning the passengers to the optimal bus and then selecting the optimal route once the assignments are made [20]. These two optimizations steps, commonly referred to as clustering and routing, can either be performed independently of one another or simultaneously [21].

The multi-bus DARP can be further divided into uniform and non-uniform systems, where a uniform system is one in which the buses all share the same characteristics (i.e., capacity, fuel efficiency, etc.) and a non-uniform system is one in which buses will have varying characteristics.

2.1.2 Online versus Offline Dial-a-Ride

The second characterization of dial-a-ride systems deals with the manner in which passengers are added to the system. The first type is an offline system in which all passenger information is gathered before optimization begins and new passengers cannot be added once the system has been optimized and set into motion. The second option is the online system, in which passengers may join, leave, or change their requirements at any moment. This type of system must update and quickly react to these changes [15]. In the offline problem, all passenger information is known *a priori*, which allows the optimization routine to find a very close approximation of the optimal solution. However, the offline method is largely impractical in real-world setting as it requires a critical mass of passengers to be waiting before service can begin and it requires passengers to make all of their requests well in advance of when they require service.

In the online problem, passengers and routes are updated continuously based on new passengers joining the system. This allows for last-minute requests to be made by the passengers but poses additional challenges to the operator. The primary challenge of implementing an online demand-responsive system is that of building a vehicle optimization and routing algorithm that can find optimal or near optimal solutions in real time. The added urgency of completing the optimization routine in real-time for the online case, has led to many researchers to resort to heuristic rather than optimal control techniques [18] [22]. Additional logistical challenges of the online dial-a-ride problem include providing efficient communication between passengers, dispatchers, and drivers in order to update all parties of changes to the schedule, and ensuring that existing passenger time-window constraints are not violated when new passengers are added to the vehicle [23].

2.1.3 Passenger Time-Windowing

Since a portion of the proposed research involves coordinating passenger handoffs between fixed and dynamic routes, a brief introduction to on-demand time windows is necessary. In a DARP with time windows (DARPTW), the passenger provides a time constraint during which service can be rendered and penalties are placed on the route selection routine for missing these time windows. Wang et al. provide a concise definition and formulation in [24]. Heuristic and approximate solutions to solving the DARPTW can be found in [18] [24] and [25].

2.1.4 Integrated Dial-a-Ride

Systems which utilize both fixed and dynamic transportation modes are referred to as integrated demand-responsive transportation systems and the problem of assigning and routing passengers in these types of systems is referred to as the integrated dial-a-ride problem (IDARP). An integrated demand-responsive transportation system was first suggested by Wilson et al. in 1976 [26]. Wilson suggested a network of DRT neighborhoods connected by a fixed-route transit route. Wilson as well as other researchers have continued to expand and refine the IDARP since then. For instance, in 1996 Liaw developed an approach to the IDARP which identifies the nearest fixed-route station and routes the passenger to this stop via DRT [27]. In 2003, Aldaihani and Dessouky developed a heuristic method of creating combined demand-responsive and fixed routes in a paratransit setting [22]. The study utilized real data to compare curb-to-curb service with an IDARP to save costs on the part of the operator while minimizing travel time experienced by the passengers. In their study, 18.6% of the trips in the study were converted from being strictly demand-responsive trips to integrated demand-responsive trips in order to save costs for the transit operators. These new integrated routes increased the passengers' trip time by 5.4% on average but lowered the cost incurred by the operator [22]. Depending on the application, this may be an acceptable compromise, and further refinement of their approach may lead to better results. In 2009, Hall et al. developed an in-depth formulation of the offline IDARP and introduced methods to reduce the solution search area. A simple IDARP case is presented in their work and solved analytically [28].

In addition to these applications, Lee and Wang [29] have developed a high-level architecture for a taxi-pooling system in which passengers who are not within walking distance to a transit station can utilize DRT to reach the nearest station. Another application of integrated DRT is discussed by Uchimura et al. in [20]. Uchimura suggests a hierarchical transportation system for the city of Seattle, WA where the highest levels of the system use high-speed rail for long-distance travel and the lowest levels of the system use shuttles or DRT to handle transportation within neighborhoods. Another type of integrated DRT is suggested by Crainic et al. in [7]. Crainic suggests a hybrid transportation system which is based on fixed routes with fixed stops and times but also is capable of deviating from the prescribed route to pick up passengers on demand. The efficacy of these types of systems is demonstrated by Horn in [30], in which Horn develops a planning procedure for passengers to choose between demand-responsive and fixed routing.

The network-inspired transportation system (NITS), which is the basis for the case studies presented in this research, is classified as an integrated demand-responsive transit system because it uses demand-responsive transit to handle the “first mile” and “last mile” of each passenger’s trip while using the fixed-route network to handle the bulk of the passenger’s journey.

2.2 Review of Demand-Responsive Comparison Methods

The previous section provided a general overview of the dial-a-ride problem and a survey of the various types of demand-responsive transportation implementations. This section dives more deeply into the current research related to objectively measuring the performance of demand-responsive transportation and comparing that performance with fixed-route transportation.

One of the two major goals of this research is to develop a method to objectively compare the performance of demand-responsive transportation systems to the performance of traditional fixed-routes systems in heterogeneous environments. An objective comparison method is required in order to determine which portions of a city are most appropriate for demand-responsive transportation and which portions of the city are best suited to fixed-route transportation. Current methods of comparing demand-responsive transportation (DRT) and fixed-route transportation (FRT) have focused on homogeneous street layouts. Methods by Diana et al. [31] and Thompson et al, [32], compare FRT and DRT in gridded street systems as well as ring-radial systems, common in European cities. Li and Quadrifoglio perform similar analysis of feeder systems [33][34]. The feeder systems analyzed by Quadriofoglio and Li divide a city into multiple feeder pools where each pool is rectangular in shape with a transit station at one end of the pool [35]. A feeder bus then traverses the length of the feeder pool, collecting passengers and dropping them at the transit station.

The techniques developed by Li, Quadrifoglio and Diana provide an excellent foundation for comparing DRT and FRT systems and determining when a DRT will outperform an FRT system. A key contribution of their research is identifying the basic costs of travel for the passenger and the transit operator [36]. The passenger costs are broken down into walking time for the passenger, time spent waiting for a vehicle, and the riding time for the passenger. These costs do not represent all the costs levied on the passenger, notably the fare is missing from this equation. Instead, these costs are intended to represent the variable costs to the customer. Theoretically, a transit fare is a flat rate, but the time it takes to complete a trip can be controlled by the operator in real time and can be minimized by optimizing the route taken by the vehicle. The transit operator's cost is commonly expressed as a function of vehicle miles traveled (VMT). Once again this is not intended to represent all the costs of operating a vehicle. It is merely an approximation of the costs that can be controlled by the driver, specifically how many miles the vehicle drives. These costs are what the dial-a-ride algorithms attempt to minimize when selecting optimal routes.

The main drawback of the existing techniques is that the cities and service areas studied are homogeneous, meaning that there is little randomness in their layout. Each system assumes a grid or ring-radial street layout [31] [34]. It is assumed that the dispersion of trips within a service area as well as the rates at which passengers request trips can be closely approximated by a closed-form probability density function [37]. These techniques are also designed to study transit systems with vehicles that arrive at unchanging intervals and with stations that are equally spaced within the service area. Since the service areas that these methods were developed to study possess these characteristics, it is practical to make these assumptions. Assuming a homogeneous environment allows the research to perform complex analysis by developing closed-form solutions to practical questions about the service area. These questions include finding the optimal number of vehicles required to service an area, determining the maximum trip-request rate that a demand-responsive service can handle, identifying effects of altering a transit schedule within an area, and ultimately determining whether or not an area is best suited to demand-responsive transit or fixed-route transit.

However, if the environment under study is more heterogeneous, these methods are less accurate in their estimates. If an area has a suburban street layout instead of a gridded street layout, if the transit schedule changes often during the day with no apparent pattern to the distribution of stops, and if passenger locations and request rates vary both spatially and temporally, then finding closed-form solutions to represent this environment is increasingly difficult. In order to provide the same analysis within these environments, a simulation-based approach is utilized. In the simulation-based approach, no assumption is made about passenger locations or trip request times, transit schedules, or street layout. Instead every passenger trip is modeled to exacting detail taking into account the actual street network from an area, passenger survey data that provides request times and locations, and accurate transit schedules that take into account the time of day and day of week. This simulation-based approach is the foundation for the comparison methods proposed in this research. The simulation-based comparison technique is outlined and studied in Chapter 5.

2.3 Building a Dial-a-Ride Simulator

Since this research uses a simulation-based approach to compare and optimize integrated demand-responsive transportation systems, a modular and comprehensive software simulator is needed. The simulator is intended to be used by transit operators to determine when, where, and how to build a demand-responsive transportation system and how to integrate it with existing fixed-route transit systems. The simulator described in the following sections is capable of simulating the various types of dial-a-ride systems described in Section 2.1 and is also be used to simulate the performance of fixed-route systems.

Simulating a large set of passenger trips within heterogeneous environments that include both demand-responsive and fixed-route transportation options, requires a number of sub-problems to be solved. The most fundamental problems that need to be solved are listed below.

1. How should a street network be represented, and how are optimal routes found within that network?
2. How should a complex transit schedule and map be represented, and how are optimal passenger itineraries found within that schedule and map?
3. How are optimal demand-responsive routes found, and how is the optimal vehicle for each trip identified?

2.3.1 Routing Vehicles in Heterogeneous Environments

Routing vehicles in a heterogeneous environment requires two sub-problems to be solved. First, a method representing the street system must be found. Second, an algorithm to find shortest paths within the street system is needed.

Representing Heterogeneous Street Systems

A major problem that must be addressed is that of representing a heterogeneous street network and routing vehicles within that network. The street network needs to accurately depict the true costs of traveling along each street. These costs may include travel time, average road speed, average road throughput, and traffic congestion [24]. Typically these conditions are encapsulated as a graph, $G = (N, E)$, where the nodes N represents street intersections as well as potential passenger pickup and drop-off locations within the system, and edges E represent the street connections between the nodes. Creating graphic representations of city streets and conditions is a large problem unto itself. Since the simulator developed for this research needs to be easily applied to a wide range of areas, graph-based representations of street networks need to be readily available. Fortunately, such graphs already exist that provide the necessary information as well as methods of updating the information if required. This research uses the Open Street Map project as the basis for representing city streets. Open Street Map is an open-source mapping project that allows for estimation of driving time and distance and allows for customization that would permit traffic conditions to be encoded as well [38].

Finding Shortest Vehicle Paths

Once a valid representation of the street network is found, a method of quickly finding shortest paths within that network is needed. Since the dial-a-ride optimizer requires dozens of potential trips to be calculated for every passenger before selecting the optimal trip, the method of finding shortest paths within a large network must be extremely efficient. Traditionally, the shortest paths between edges of these nodes is found via Dijkstra's algorithm, A* or similar path optimization algorithms [20] [39] [40]. These route selection algorithms encounter two difficult problems when applied to large street network system. The first problem is their inability to scale with large networks. For instance, Dijkstra's algorithm is of $O(n^2)$ complexity where n is the number of nodes in G [24]. The second problem is that these systems need to be able to react to changes in the system and propagate this information throughout the system very quickly. The standard Dijkstra's algorithm may not be able to update quickly enough if new road conditions are continuously updated.

To work around these issues, several improvements in the conventional route selection algorithms have been proposed in previous research. Notably these methods include the double bucket approach taken by Eklund et al. and Wang et al. [24]. The double bucket approach, described in [39], ensures that each node is scanned at most once when executing Dijkstra's algorithm. Kim et al. utilize a state space reduction technique which utilizes two-state Markov chains to represent the congestion status of each link. The specifics of this approach are described in [41]. Another approach by Miller-Hooks and Mahmassani associates a probability density function (PDF) of the cost of each link and then takes the expected value of this PDF to represent the cost of the link [42]. At this point the problem can be solved in a stationary manner.

In order to find routes in real-time, computation time must be kept low. For a continent-wide street network, this is a challenging problem. However, computational speed issues have largely been addressed through hierarchical routing. Hierarchical routing methods, such as highway-node routing, separate the street graph into hierarchies where large and high-speed highways are at the top of the hierarchy and small local streets are at the bottom of the hierarchy with various feeder streets and boulevards occupying the space between. The goal of highway-node routing is to use the high speed highways for the bulk of each journey. In this way many possible routes are removed from the search space and the numerous local streets only need to be considered for the first and last portion of each trip [43] [44]. The computation time can be further reduced by pre-calculating travel times and paths between popular nodes. This technique is known as transit-node routing [43] [45].

The previous routing techniques focus mainly on selecting optimal routes for single vehicles. Even those researchers who are attempting to encode the network for use in a DRT scheme often overlook a common problem which only occurs in routing vehicles with intermediate stops; i.e., handling U-turns. When routing vehicles between only two stops, the U-turn is rarely necessary. However, when routing a vehicle between many stops, performing a U-turn may be necessary or when U-turns are not possible, a workaround must be chosen. Adapting the conventional street network graph to handle U-turns is addressed by Fan et al. by adding weights to the nodes as well as the edges [46].

Practically speaking, developing a routing engine is an expensive endeavor in terms of time and resources. Fortunately, open-source routing and mapping software is available. The Open Source Routing Machine (OSRM) [47] is an open-source routing engine that was developed by Peter Sanders et al., as a practical implementation of many of the routing algorithms discussed above. Most notably, the OSRM takes advantage of hierarchical routing algorithms to quickly find optimal routes in large networks [43]. The OSRM is continuously updated with advances in routing, featuring accurate and fast hierarchical routing and turn-restricted routing that is necessary for trips with intermediate destinations. Furthermore, since the Open Source Routing Machine utilizes Open Street Map graphs, OSRM and Open Street Map provide a foundation of software on which innovative demand-responsive routing techniques can be tested.

2.3.2 Finding Optimal Transit Itineraries

In addition to routing vehicles along the road network, a method of routing passengers within the transit network is also needed. This information is needed for two reasons, for comparison between fixed-route and demand-responsive travel and for routing passengers via fixed-route services in an integrated transportation system. As with vehicle routing, a method of representing the transit schedule and map is needed. The most popular representation of transit schedules and maps is the general transit feed specification (GTFS) [48]. GTFS encodes both the schedules and maps of transit agencies and is widely available for a large number of agencies [49].

Transit routing algorithms extend traditional graph-based searches with the additional constraint that the presence of edges follows a schedule. For example, while a road will always be present in a road network, a bus or train is only present for short intermittent times. To account for the intermittent presence of edges, schedule-based path planning algorithms must be created. Schedule-based routines are discussed in [50] [51].

This research uses the open-source project known as OpenTripPlanner for transit routing purposes. The OpenTripPlanner, originally developed for TriMet in Portland, OR, utilizes the General Transit Feed Specification as well as Open Street Map street data for optimizing point-to-point transit trips. The routing algorithm is built upon the schedule-based routing algorithm of [51].

2.3.3 Solving the Dial-a-Ride Problem

Once a foundation for finding optimal vehicle routes between two points is found, the next problem is finding optimal routes between many points. This is the dial-a-ride problem discussed earlier in this chapter. The dial-a-ride problem is a modified version of the traveling salesman problem (TSP) and approaches to solving the dial-a-ride problem have been heavily researched. Like the traveling salesman problem, the dial-a-ride problem is NP-hard and becomes exponentially more difficult to solve as the number of busses and passengers increases in the system [52]. One factor leading to the difficulty in solving the traveling salesman and dial-a-ride problems is the larger number of possible permutations. For *N* cities in the traveling salesman problem, there are $N!$ possible permutations of the order in which to visit those cities. Like the traveling salesman problem, the dial-a-ride problem also suffers from this scalability issue. The DARP seeks to select the order in which to visit a pick-up and drop-off location for *N* passengers. For *N* passengers, there are 2*N* locations to be visited. Consider

$$R = [P_{1,P}, P_{1,D}, P_{2,P}, P_{2,D}, \dots, P_{N,P}, P_{N,D}],$$

where *R* is the set of all passenger locations and $P_{i,P}$ and $P_{i,D}$ are the pick-up and drop-off locations of passenger i, respectively. Since there are 2*N* locations in *R*, there are (2*N*)! ways in which to visit the locations. However, some of these permutations are not valid. These invalid permutations are those in which a passenger's drop-off location is visited before his or her pickup-location. Dividing the total number of permutations by the number of ways in which illegal visits occur, yields the following formula for counting possible permutations,

$$\frac{(2N)!}{2^N}.$$

This means that for a DARP consisting of 5 passengers, there are 113,400 possible permutations to consider For 10 passengers this number jumps to $2.38\text{x}10^{15}$. For a city-wide demand-responsive transportation system, there could be hundreds or even thousands of passengers requesting service at any given time. Determining the optimal route under these circumstances is infeasible. Because of this, analytical solutions are rarely found outside of very simple instances of the dial-a-ride problem.

Popular methods of solving the DARP include numeric estimation methods such as genetic algorithms (GA) [24] [20] [21], tabu search [53], or simulated annealing [52] where the order in which passenger locations are continually adapted and improved until a near-optimal solution is found. A thorough review of various optimization methods for the dial-a-ride problem can be found in [15]. While these methods will find near-optimal solutions, they are also time-consuming for large instances. For problems consisting of 250-300 passengers, these search methods can take between 10-90 minutes depending on the accuracy requirements of the particular application [21]. For the offline DARP, when trip requests are made hours or even days in advance, this computation time is not critical. However, in a real-time scenario, this time delay is not acceptable.

In a large city, dozens of passengers may be requesting trips each second, and the dial-a-ride solver must be able to find near-optimal solutions real time [54]. One approach to finding near-optimal solutions is to reduce the size of the search space. Branch and cut and heuristic search space reduction techniques are presented by Kim et al. and Cordeau in [41] [54] with promising results, however several minutes are still required to find a solutions for problems with more than 100 passengers. Faster methods of solving the DARP are rule-based insertion algorithms. For the online version of the DARP, Jaw et al., Aldaihani and Dessouky, and Miyamoto et al. have all proposed methods for quickly inserting new passengers into the system [18] [22] [25]. As each new passenger is inserted into the system, the order of passenger locations previously identified will not change. Instead, the new passenger pickup and drop-off locations will be inserted into the existing order. This will cause the visit times for some locations to change, but the order will be maintained.

Insertion heuristics drastically reduce the search space but do not necessarily reduce the performance of the system by a large degree. For instance, Aldaihani and Dessouky compare a simple heuristic insertion problem with a more comprehensive tabu search method. Their heuristic method found a solution to a medium-sized DARP (30 passengers) that required the vehicles to travel 6% farther than a tabu search solution with passengers spending, on average, 8% more time on the vehicles. However, the tabu search required 623 seconds to complete the optimization routine while the heuristic method only took 9 seconds, illustrating the potential for heuristic methods in the dynamic DARP. The optimization method used in this research, which is described in detail in Chapter 4, is a sequential heuristic insertion algorithm.

2.4 Optimally Designing Demand-Responsive Feeder Zones

Much of the work discussed in the previous sections of this literature review deals with optimizing routes and vehicle selection within a predefined service area [28] [30] [22] [55]. However, as this research will show, the shape and size of the demand-responsive service areas themselves can have drastic impacts on the performance of the system as a whole [33]. For instance, larger demand-responsive feeder zones require the demand-responsive vehicles to travel longer distances and passenger wait times can be increased due to this. On the other hand, smaller demand-responsive zones can reduce the vehicles miles traveled and keep passenger wait and ride times low, but the tradeoff is that fewer trips will be eligible for demand-responsive service.

A major goal of this research is to develop a method of selecting the optimal shape and optimal size of a demand-responsive zone in heterogeneous conditions. Heterogeneous conditions refer to street networks that do not follow a strict grid pattern, passenger arrival rates that are not constant, passenger locations that are not evenly spread throughout a region, and transit schedules that that are not consistent throughout the day and do not have a regular pattern of stops. The optimal design of a feeder zone is one that maximizes customer satisfaction, in terms of minimizing waiting and riding times, minimizes operator costs in terms of vehicle miles traveled and reaches the highest number of passengers within the system.

Suggestions for designing fixed-route transit systems in heterogeneous environments are provided by Chien and Schonfield in [56]. Chien and Schonfield propose a method of optimizing fixed-route transit that does not make system-level assumptions about transit arrival rates, passenger distribution, or passenger arrival rates. They mitigate the need to use generic passenger transit data by dividing a given service area into smaller sub-regions. Within each sub-region the passenger and transit data can take on a homogenous form without making large-scale assumptions about passenger data throughout the system. Instead they use travel demand modeling to estimate the passenger departure and arrival rates within smaller service areas where travel is likely to be more homogenous.

In work by Chang and Schonfield [57], the optimal number of service zones within a rectangular service area is found. In addition to the optimal number of zones, optimal vehicle size, optimal headway, optimal number of vehicles, and optimal routes are found within each zone. While homogeneity is not assumed across the entire service area, within each zone, homogeneity is assumed in regards to passenger distribution and traffic conditions, e.g., speed. Therefore, unless each zone is kept small, which would affect performance of the system, accuracy will be lost in optimizing routes within each zone. The research presented in this dissertation seeks to provide granular vehicle and route optimization across a large service area without assuming homogeneity at any level.

Quadrifoglio, Li, Aldaihani, et al., have provided significant contributions towards designing optimal demand-responsive feeder systems in recent years. Their work provides an excellent framework in which to design optimal feeder zones and focuses on designing optimally-sized feeder zones for hybrid grid systems. Aldaihani et al., in [58] developed a method to determine the optimal layout of transit and demand-responsive zones in a gridded street system. This basic layout for this system is shown in Figure 2-1. The system is said to be square and the size and location of the system is given. The objective of Aldaihani's work is to optimally select the number of zones within this system. The square service area will be evenly divided into nxn service areas connected by a fixed-route bus line at the center of each zone. If n=1, then there is only one zone and no need for a fixed-route system. All trips within the service area will be serviced via demand-responsive vehicle. In this study, the demand-responsive vehicles act like a taxi service. They provide point-to-point travel for each passenger but ride sharing is not allowed. Only one trip per vehicle is permitted. As n increases, the size of each feeder zone decreases and the probability that a passenger will require a trip outside the zone also increases.

Methods

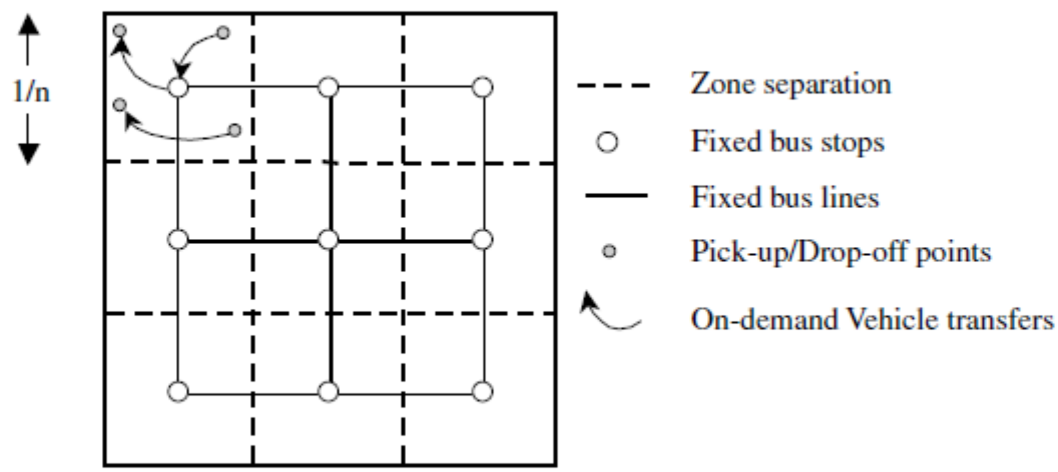

Figure 2-1: Design of a hybrid feeder system with 3x3 zones [58].

Selecting the optimal n will trade off between passenger costs and operator costs. As n increases, the probability that a passenger will require a transfer between fixed-route and demand-responsive transportation also increases, increasing passenger travel times. However, for the transit operator, it is less costly to utilize fixed-route transit than demand-responsive transit in this scenario, because fixed-route transit allows for ride sharing. Conversely, as n decreases, passenger times will decrease, but operator costs will increase. Given this scenario, closed-form solutions can be found analytically to represent the passenger travel and wait times as well as the transit operator costs. The costs measured in this work combine the cost of operating flexible-route vehicles, fixed-route vehicles, as well as the perceived costs incurred by the passenger. The costs are all converted into monetary values representing expected costs for vehicles and passenger time. For instance, flexible route costs consist of the cost of owning, operating, and maintaining a flexible route vehicle per day as well as additional costs incurred by increased usage during the day. Passenger costs are assigned a monetary value based on the passenger's time. Passenger costs are not necessarily uniform across all modes as waiting and riding times are not perceived equally for all modes [59].

In another type of optimization format, Xiuguang Li et al., [60] [33], developed a method to determine the optimal size and optimal number of feeder zones in a rectangular feeder system connected on one end to a major transit network. Figure 2-2 shows the basic layout of this system. The zones shown in Figure 2-2 act as origins and destinations for passengers. Passenger leaving the system will be transported from their origin location to the transit terminal located at the far end of their zone. Fixed-route transit will then handle the remainder of the journey. Similarly, a passenger returning to one of the zones will be transported from the terminal to his or her final destination within the zone. Transportation within the zone is handled with either demand-responsive or fixed route transportation depending on the passenger demand at the time of the request. The objective of Li's work is to determine the optimal width (W) and length (L) of the zones as well as to determine the optimal number of feeder zones within that area. Li extends this work by considering the presence of two vehicles within each zone [37]. The presence of additional vehicles allows the zones to be further divided into more sections which increases the complexity of the problem but also allows for improved performance.

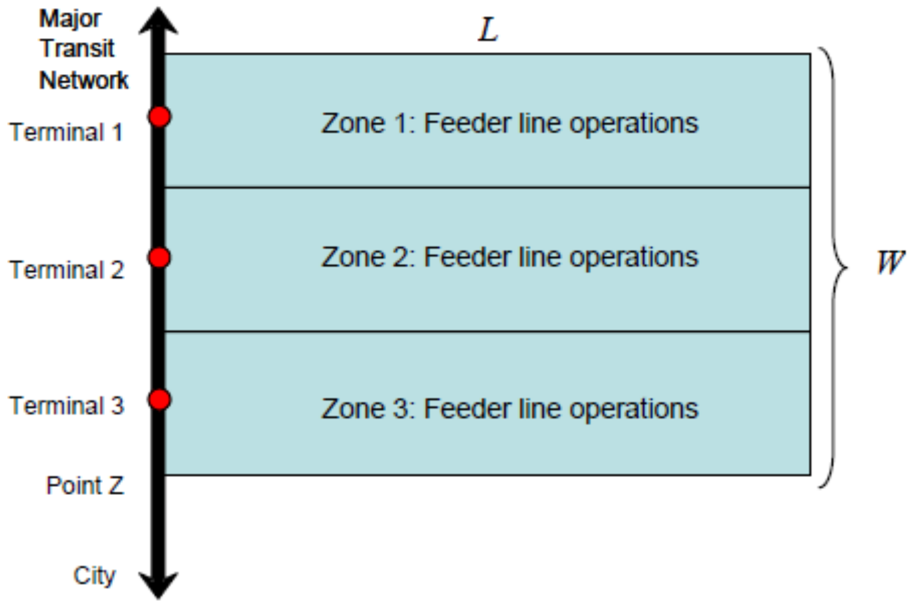

Figure 2-2: Rectangular feeder system with 3 transit terminals.

The work by Li leads to the creation of closed-form analytical solutions to complex transportation optimization problem, and as with Aldaihani's work, the costs for the transit operator and passengers are converted into easily compared monetary values representing the cost of operating the vehicles and the value of the passengers' time. However, a drawback to this approach is that granular transit schedule and passenger information, such as that found in the work of Chien and Schonfield, is not taken into account. Assumptions are made about the passenger request rate and location, specifically passenger request rates are assumed to be uniform in time and passengers are assumed to be spread evenly across the service area. Also, the street system is assumed to be a grid. Street systems that do not approximate a grid or another regular form will be difficult to model using this approach.

The design methods developed by Li and Quadrifoglio provide a strong framework within which to determine the characteristics of optimal demand-responsive feeder zones. Importantly, they identify the major costs associated with demand-responsive service, both for the passengers and the transit operator. These costs, which include passenger walk, wait, and ride time as well as vehicle operation and maintenance costs, are reduced to a monetary representation allowing for the total passenger and operator costs to be optimized. A breakdown of these costs is given in [33].

One of the original applications of their methods was to design a feeder system for El Cenizo, TX, shown in Figure 2-3. The regular grid layout of this city, the presence of a major transit station along the edge of the zone, and the ability to represent passenger dispersion and ride request rates with closed-form expressions make this a relatively homogeneous environment. Therefore, the analytical solutions found from their research provide an accurate and effective means of identifying optimal feeder zone design in this area.

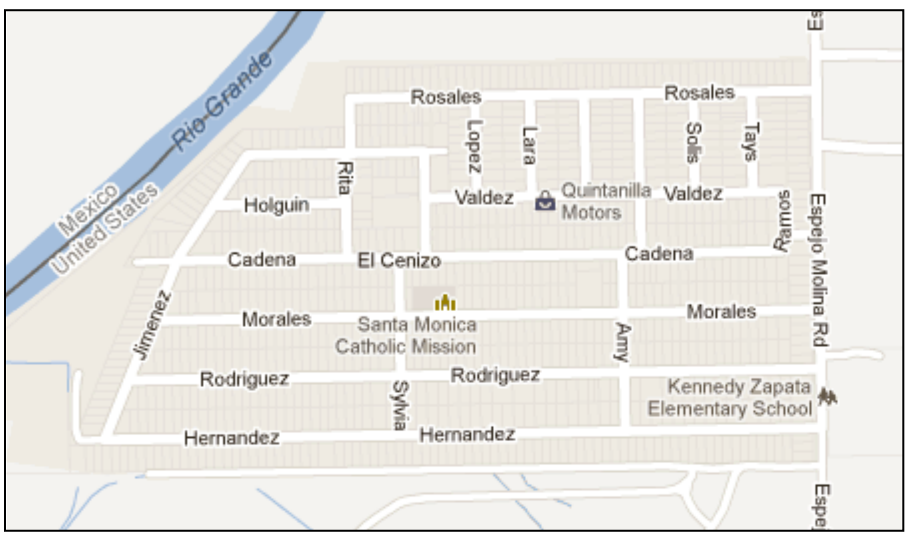

Figure 2-3: Map of El Cenizo, TX demand-responsive feeder areas.

Another study done by Quadrifoglio, Dessouky, and Ordonez in [61] seeks to identify the zoning strategy and time-windowing policy for a demand-responsive transportation system in Los Angeles County. The demand-responsive system under study was a paratransit service aimed at mobility impaired passengers across six large service regions within Los Angeles. In this study, the shapes and layouts of the regions as well as the dispersion of passengers within the regions are highly non-uniform. Because of this, a simulation-based approach was taken to optimizing the design parameters of time-window size and selecting a transfer policy between the zones. In the simulation, the locations of passengers are assumed to be known ahead of time and an offline dial-a-ride problem is then solved using a sequential heuristic approach. Many simulations were run comparing various time window sizes and zone policies. In the study, passenger time windows were varied from 10 to 45 minutes. The study showed a linear relationship between time window size and demand-responsive system performance. In terms of operator costs, larger time windows results in fewer miles traveled by the fleet and fewer vehicles being needed to meet demand. However, larger time windows resulted in longer waits and ride distances for the passengers.

The second optimization variable was that of selecting a zone policy. Four of the six zones were combined into a variety of different super zones to study the effects of combining zones. At one extreme, each zone acted independently and handled all trips originating within the zone without outside assistance. In the other extreme, all zones were combined into one large zone and vehicles were permitted to travel anywhere within the region. Other policy choices involved combining two or more zones into larger regions while other zones remained independent. For each policy choice, simulations were run to determine the average operator costs and passenger costs. In the particular study conducted for Los Angeles, one large super region was found to be the most cost-effective for the operator and passengers. Since this study assumed that ride requests were submitted one day in advance, the large processing time necessary to route passengers across a large area is not an issue.

There are two important aspects learned from this research. The first is the relationship between time windows, passenger satisfaction, and vehicle operating costs.

As time windows are allowed to increase, the number of vehicles required to meet passenger demand decreased. In the study, each minute added to passenger time windows resulted in 2 fewer vehicles used and 260 fewer miles driven. However, wait times and dependability from the passenger's perspective suffered, indicating a tradeoff between operation costs and passenger satisfaction that can be affected by adjusting the passenger time windows. The second aspect is the notion of using simulation-based approaches to assess changes in geographic areas. In the work of Quadrifoglio et al., six regions were predefined and various combinations of these regions were studied. This dissertation takes a similar simulation approach that identifies the optimal shape and size of the zones without requiring layouts being defined *a priori*.

In other zone optimization research, Pages et al. [62], developed a method of solving very large scale routing problems in real time. Their method includes a step that groups passengers into zones before determining the optimal routes of the vehicles traveling within and between those zones. The novelty of their method is that instead of looking at each individual trip and optimizing it, they examine flows of passengers between zones. Their method vastly reduces the amount of time it takes to solve a vehicle routing problem for large scale systems by dividing the system into smaller zones. Since these zones are not created arbitrarily, but are based on expected passenger flows between the zones, global cost minimization is considered in their creation. While this method provides a fast way of optimizing the vehicle routing problem, it doesn't fully address the needs of this research. This research presented in this thesis seeks to define the line between one or more demand-responsive service areas as well as the boundary between fixed-route and demand-responsive service areas. This is in contrast to the work of Pages where all are assumed to be within a demand-responsive service area.

The zone optimization research discussed in this literature review often simplifies the problem to such a degree that those optimization techniques cannot be applied to heterogeneous environments. In research where heterogeneous environments are analyzed, the shape or size of the demand-responsive zones is already determined, and the focus of the optimization becomes selecting other characteristics, such as the number of vehicles or the optimal transfer policy between zones. This research will suggest a zone design approach that will identify the optimal shape and size of demand responsive feeder zones for areas with disorderly street patters, irregular passenger behavior, and complex transit schedules. Due to the irregular nature of the problem, a simulation-based approach will be used that takes advantage of the integrated demand-responsive transportation simulator described in Chapter 4. The optimization approach will be highly generic and applicable to a wide range of environments and is a major contribution of this research. The zone design optimization method proposed in the research is discussed in detail in Chapter 6.

CHAPTER 3

THE NETWORK-INSPIRED TRANSPORTATION SYSTEM

The previous chapter provided background on various types of demand-responsive transportation systems. The demand-responsive systems that are analyzed in this research are part of a proposed integrated demand-responsive transportation system framework known as the network-inspired transportation system (NITS). In this chapter, the NITS framework is defined and the basic goals of NITS operation are outlined. The NITS framework serves as the foundation for the demand-responsive simulator that was built for this research in order to compare and optimize demand-responsive transportation networks in heterogeneous environments. The NITS will act as the framework for the case studies performed in Chapters 5 and 6. Details of the simulator are presented in Chapter 4.

3.1 Description of the NITS

The network-inspired transportation system (NITS) is a framework in which passengers are transported through a complex metropolitan transportation system in the same manner that data packets are routed through a telecommunications network. In the NITS framework, transit infrastructure is treated as a packet-switched network where passengers are analogous to data packets, high-speed rail and bus rapid transit (BRT) are analogous to high-speed data trunk lines, and areas of demand-responsive service are analogous to subnetworks or subnets.

A packet-switched network is highly hierarachical. If a data packet is routed between two computers in a local area network, the packet will never leave the local area network. However, if a data packet is destined for a server that is in a different building, city, or state than the origin, then the local area network will not handle end-to-end routing. Instead, the local area network will recognize that the packet is destined for a server outside the network and will simply pass the packet to the gateway. Once handed off to the gateway router, the path of the packet is irrelevant to the original local area network. The packet will continue to pass through a hierarchy of routers and networks until it reaches the final destination. Packet routing allows individual networks to operate semi-independently of one another yet permits them to cooperate in order to route data, very efficiently, across large distances.

The same type of hierarchical routing used for data networks can be used to route passengers in a complex transit system. If a passenger wishes to makes a short trip within a neighborhood, a single vehicle will handle the entire trip door-to-door. However, if a passenger wishes to travel across the city, multiple legs will be needed to complete the journey. The first leg of the journey will be handled via a demand-responsive vehicle. The driver of this vehicle will know that this passenger's destination is not within his or her coverage area, so the driver will not be handling the entire journey for that passenger. Instead the driver will route the passenger to a transit station within the coverage area. The transit station acts like a gateway between the demand-responsive zone and the fixed-route network. Once in the fixed-route network, the passenger may take one or more trains or buses to get to a transit station near his or her final destination. Like the first leg of the journey, the final leg will also be handled with a demand-responsive vehicle. The passenger will be transported from a transit station within the demand-responsive coverage area of his or her destination to the final destination. In this manner, demand-responsive transportation handles the first-mile and last-mile of the journey and high-speed, fixed-route transit handles the bulk of the journey.

The network-inspired transportation system consists of a set of demand-responsive zones or subnets as well as a fixed-route network. A demand-responsive subnet is defined as an area in which a passenger can be routed via a demand-responsive vehicle without the use of fixed-route transit. Each demand-responsive subnet acts as a semi-independent demand-responsive system. Typically, a subnet will consist of one or more adjacent neighborhoods. Selecting the optimal size and layout of the demand-responsive subnets is a primary objective of this research and is throughly discussed in Chapter 6. The two main roles of the subnets are to provide demand-responsive transportion to and from local transit stations as well as transportation within the subnet. In order to provide service across an entire city, it is possible to keep the individual subnets small while using a fixed-route system of transit to connect the subnets. The advantage of keeping the subnets small is that the computation time required to solve the dial-a-ride problem within each subnet is much lower than attempting to solve a city-wide dial-a-ride problem.

The fixed-route network can consist of bus rapid transit (BRT), rail, and local bus service. These modes do not have the ability to alter their routes on-demand, but they offer other advantages in terms of operating costs. The cost per vehicle associated with fixed routes is largely independent of the number of passengers currently utilizing those modes. For instance, transporting 10 passengers along a bus route is not much more expensive than transporting 15 passengers on the same bus route. Therefore, when possible and practical, on-demand vehicles should transfer passengers to the fixed-route network when it reduces global operating costs.

Besides potential operator cost savings, BRT and rail offer benefits to the passengers, such as high speed travel and infrequent stopping due to their dedicated right-of-way. For this reason, rail and BRT can provide the high-speed, high-capacity backbone for the transit system. Using the analogy of the telecommunications network, these fixed-route lines are like high speed trunks that connect smaller subnetworks across potentially large physical distances.

It is worth noting that the nature of these subnets and fixed-route transit networks is highly scalable. Much as packets routed through the internet are passed up a chain of larger and larger networks, passengers wishing to travel between metropolitan areas can be passed through layers of larger and larger transit networks. For instance, a regional network can be made of many metropolitan networks serviced either by high-speed inter-city rail, bus, or even aircraft. A passenger wishing to travel between two cities can take demand-responsive transportation to the nearest local metro station. A metro train can be taken to an inter-city rail or bus depot. Inter-city rail can take the passenger to a transit station in a distant city where the passenger will then be passed down through a hierarchy of transit modes before finally using demand-responsive transit along the last mile of the jouney.

3.2 NITS Objective Functions

The objective of the NITS is to provide a framework that minizimizes the computational overhead required to determine optimal routes for passengers and vehicles in a complex transit system. This is achieved by reducing the search space. Instead of searching for optimal demand-responsive routes across an entire city, only routes within subnets are considered. The problem is reduced to minimizing operator and passenger costs within the demand-responsive subnets. Operator costs can typically be simplified to a combination of the distance traveled by each vehicle and the incremental cost of travel for each vehicle. Passenger costs are often associated with time, typically walking time, waiting time, and transit time.

The general global objective function for the NITS is shown in Equation 3.1, where α and β are the weights associated with operation costs (J_O) and passenger costs (J_P) respectively. Minimizing this cost function is the ultimate goal of this research (i.e., improve the traveler's experience while also minimizing costs incurred by transportation operator). The exact values of these weights and compositions of passenger and operator costs can vary widely depending on the goals of the operator, implementation-specific costs, and individual passenger demands. However, some assumptions and suggestions can be made.

$$J_{Total} = \alpha J_O + \beta J_P \tag{3.1}$$

First consider the operator's side of the cost function. The operator may be responsible for operating on-demand vehicles as well as fixed-route vehicles. The total cost of operating each of these types of vehicles is represented by J_D and J_S respectively.

$$J_O = \alpha_D J_D + \alpha_S J_S \tag{3.2}$$

For the purposes of this research, the costs of operating fixed-route services are fixed, and therefore, can be ignored. These means that the operator's component of the objective function can be reduced to Equation 3.3, where d_i is the cost associated with operating dynamic vehicle i and *M* is the total number of on-demand vehicles in the network.

$$J_D = \sum_{i=1}^{M} \alpha_i d_i \tag{3.3}$$

The passenger costs are shown in Equation 3.4, where p_i is the cost of routing passenger i.

$$J_P = \sum_{i=1}^{N} \beta_i p_i \tag{3.4}$$

Passenger costs can result from a number of sources depending on the demands of the individual passenger. These costs can include door-to-door time, time waiting for the bus, total time spent in transit, transfer time, and others. The effect of this is that the optimal route selected for a vehicle within a given subnet will depend on factors outside of that subnet. For instance, when attempting to minimize the door-to-door time for a passenger, the fixed-route transit schedule must be considered when optimizing the demand-responsive vehicle route. It provides no benefit to the passenger if he or she is deposited at a transit station when no transit vehicles will be arriving for a long period of time.

The framework and objective functions that define the network-inspired transportation system offer a generic model on which complex demand-responsive transportation systems can be built. This generic framework is the foundation for each of the case studies in this research. Both simple studies, such as comparing performance in a gridded street system, as well as more complicated studies, such as optimizing the size and shape of subnets in heterogeneous street networks, are modeled as network-inspired transportation systems. In order to properly study the perfomance of the network-inspired transportation system, and to optimize the characteristics of the subnets, a software simulator was developed. This software simulator is thoroughly described in the next chapter and case studies built upon this simulator will be presented in Chapters 5 and 6.

CHAPTER 4

SIMULATING DEMAND-RESPONSIVE AND FIXED-ROUTE TRANSPORTATION SYSTEMS

In order to perform comparisons between demand-responsive and fixed-route transportation as well as to optimize the size and shape of demand-responsive feeder systems in heterogeneous environments, a modular transportation simulator must be available. This simulator should be able to simulate passenger trip requests that are derived from survey data or origin-destination models and identify optimal fixed-route trips that will serve those requests. In addition to identifying fixed-route trips, the simulator should be able to simulate demand-responsive vehicle behavior in purely demand-responsive systems as well as simulate demand-responsive vehicles that interact with fixed-route systems in feeder-zone environments.

A review of available software and optimization techniques reveals that pieces of this complex simulator have been built. For instance, fixed-route schedule data and fixed-route trip planners are available in the form of the general transit feed specification (GTFS) and various fixed-route optimizers such as Google Transit, OpenTripPlanner, or Hop Stop. Point-to-point vehicle route optimizers that take advantage of open-source geographic information systems are also available, such as Open Street Map and the Open Source Routing Machine. However, in order to simulate passenger itineraries that combine fixed-route transportation with demand-responsive transportation, a multi-modal transportation simulator is needed. The simulator needs to handle more than fixed-route and point-to-point vehicle optimization independently. The simulator needs to be able to find solutions to multi-vehicle dial-a-ride problems that interact with the fixed-route network. Extensive reviews of available software revealed that no transportation simulator is readily available to meet this need.

To satisfy the lack of such a simulator, a significant programming effort was invested to build a highly customizable demand-responsive simulator. The simulator, referred to as the network-inspired transportation system (NITS) simulator, leverages existing fixed-route transit optimizers, open-source mapping software, and open-source point-to-point vehicle routing software to create a modular transportation simulator than can handle simulating individual passenger trips on fixed-route, demand-responsive, and integrated demand-responsive trips. This simulator is the primary tool used by this research to compare performance of various transit systems and to optimize the size and shape of demand-responsive feeder zones.

4.1 Overview

The demand-responsive transportation simulator used in this research is a combination of existing open-source software projects and custom-built software that is used to model the various aspects of fixed-route and demand-responsive transit. The simulator creates optimal fixed-route transit trips and optimal demand-responsive transit trips within a predefined zone.

Fixed-route transit trips are optimized using an open-source software platform known as the OpenTripPlanner. Point-to-point vehicle route optimization is handled by the Open Source Routing Machine software project. On top of these two routing platforms, custom software was created to handle the interaction between fixed-route vehicles and demand-responsive vehicles as well as to track each passenger's individual itinerary within the system.

4.2 Types of Trips Handled by the NITS Simulator

The NITS simulator is intended to handle a variety of demand-responsive and fixed-route transit environments. Within these environments various types of trips must be simulated. At the highest level, these trip types include fully fixed-route trips, fully demand-responsive trips, and hybrid trips consisting of both fixed-route portions and demand-responsive portions. The following section will describe each type of trip and explain how the NITS simulator optimizes passenger itineraries and demand-responsive vehicle routes for each type of trip.

4.2.1 Fully Fixed-Route Transit Trips

In a fully fixed-route trip, no demand responsive vehicle is utilized. Figure 4-1 shows the basic data flow within the simulator for this trip. When a passenger requests a trip, the passenger's start and end locations are identified and a fixed-route trip planner is used to find the optimal transit trip between those two points. In the NITS simulator, OpenTripPlanner is used to identify optimal transit trips. More information on the OpenTripPlanner tool is available in Section 4.3.1. Once each trip is optimized, passenger travel times and itineraries are stored for later analysis. The fixed-route trip is the simplest of all passenger trip types because the optimal route chosen for each passenger does not affect the itineraries of other passenger, making optimization of a single passenger trip independent of other passenger trips. This independent nature is not necessarily true for demand-responsive trips.

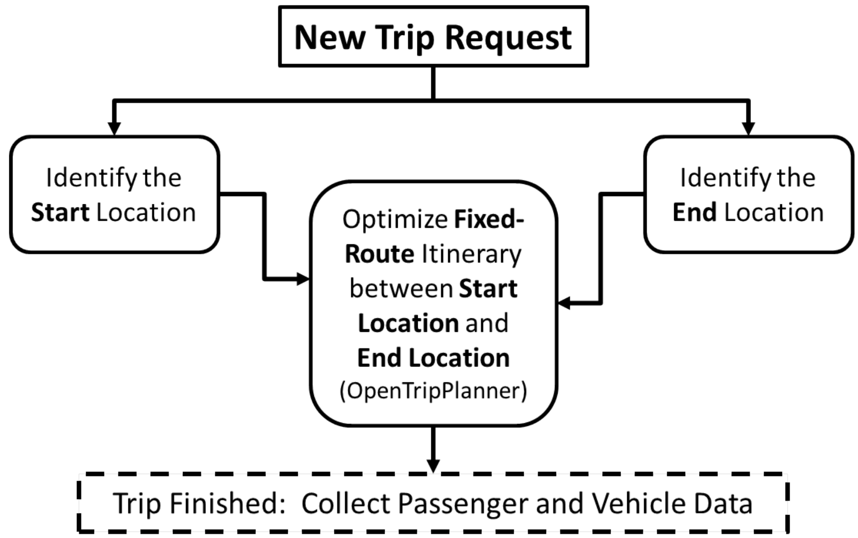

Figure 4-1: Data flow of a fixed-route transit trip.

4.2.2 Fully Demand-Responsive Trips

In a fully demand-responsive trip, the entire trip is handled with a demand-responsive vehicle. No fixed-route transit is utilized. These types of trips occur when a passenger is traveling within a single demand-responsive zone. The data flow for a fully demand-responsive trip in the NITS simulator is shown in Figure 4-2. When a fully demand responsive trip is requested, the passenger's starting and ending locations are identified and a three-step optimization routine is run to identify the passenger's optimal itinerary. The three steps include identifying the passenger's demand-responsive zone, assigning the passenger to the optimal demand-responsive vehicle, and optimally routing that vehicle. Unlike the fully fixed-route trips, when passengers request demand-responsive trips, the itineraries of other passengers are affected, i.e., when a passenger requests a trip on a demand-responsive vehicle, the other passengers on that vehicle as well as the other passenger waiting to be serviced by that vehicle will be affected. The effects felt by the existing passengers must be considered when selecting routes for the new passenger. Sections 4.3.1 – 4.3.6 will provide additional detail of how these three optimization steps occur as well as how the itineraries of other passengers are considered.

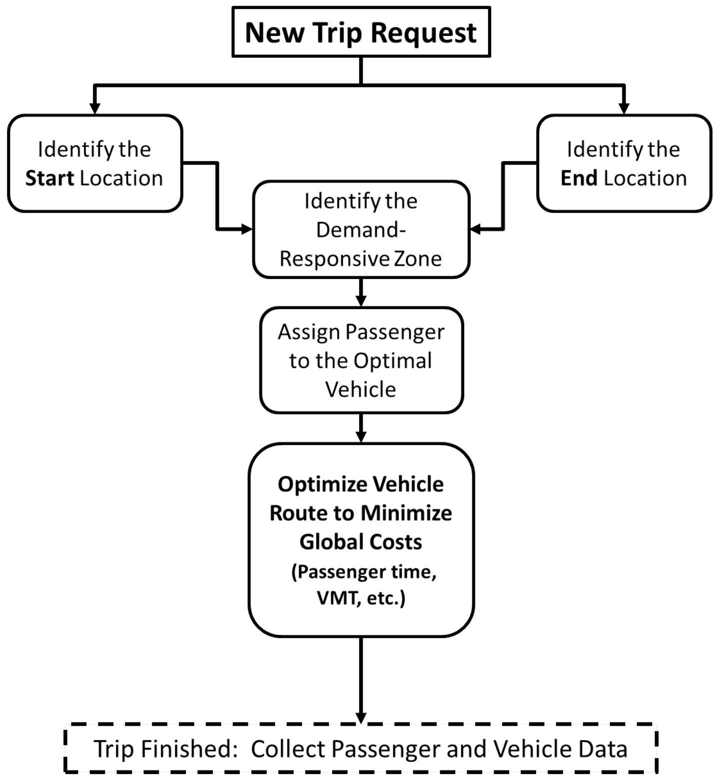

Figure 4-2: Data flow of a demand-responsive transit trip.

4.2.3 Hybrid Trips

In a hybrid trips, passengers utilize both fixed-route and demand-responsive transit to travel between their starting and ending locations. Hybrid trips can be broken down into three variations: trips that begin and end using demand-responsive transportation with a fixed-route leg in the middle, trips that begin on a demand-responsive vehicle but end on a fixed-route vehicle, and trips that begin on a fixed-route vehicle and end on a demand-responsive vehicle. Generally speaking, the demand-responsive portions of each of these types of trips are handling the first or last miles of the passenger's trip while the fixed-route portion of the trip handles the majority of the trip utilizing rail or bus rapid transit.

4.2.3.1 Demand-Responsive to Fixed-Route to Demand-Responsive Trips

In a demand-responsive to fixed-route to demand-responsive (DFD) trips, the first and last miles of the passenger's trip are handled via a demand-responsive vehicle, and the middle leg of the journey is handled via fixed-route transit. In a typical DFD trip, the first leg of the journey will take the passenger from his or her origin to a nearby gateway. The second leg will transfer the passenger between two distant gateways via fixed-route transit, and the final leg will consist of moving the passenger between a gateway near his or her destination to his or her final destination.

The basic data flow is shown in Figure 4-3. For each passenger request, the start and origin destinations are identified. For the starting location, the optimal starting zone, gateway, demand-responsive vehicle, and route are identified to route the passenger between his or her origin and a nearby transit gateway. For the destination location, the optimal feeder zone and gateway are identified. At this point, an optimal fixed-route itinerary between the origin gateway and destination gateway is found using OpenTripPlanner. Once the time-of-arrival at the destination gateway is known, an optimal demand-responsive vehicle and demand-responsive vehicle route is chosen to route the passenger between the destination gateway and the passenger's final destination.

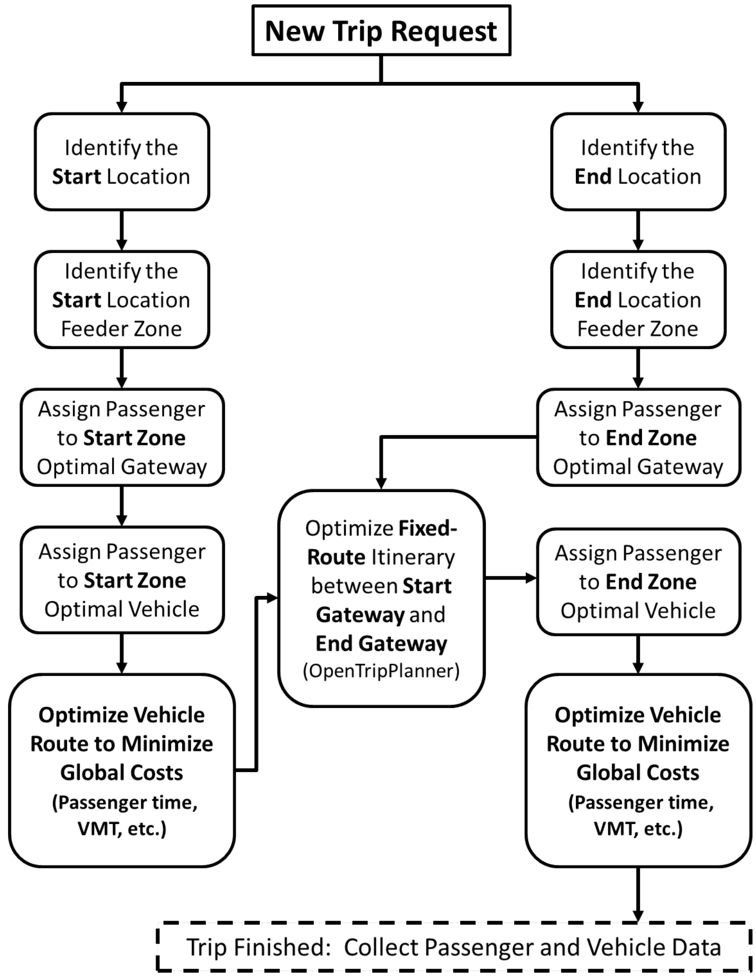

Figure 4-3: Data flow of a trip which uses demand-responsive transportation for the first and last legs and fixed-route transportation for the middle leg of the trip.

4.2.3.2 Demand-Responsive to Fixed-Route Trips

In a demand-responsive to fixed-route (DF) trip, the first leg of a passenger's journey is handled with a demand-responsive vehicle, and the remaining portion of the journey is handled with fixed-route transit. The role of the demand-responsive vehicle is to transport the passenger between his or her origin location and a nearby optimal transit gateway. The data flow of a DF trip is shown in Figure 4-4. As with DFD trips, the passenger's origin location is identified and a series of optimization steps are undertaken to route the passenger between his or her origin location and a transit gateway. Upon reaching the gateway, the remaining portion of the trip will be handled via fixed-route transit and will be optimized using the OpenTripPlanner tool.

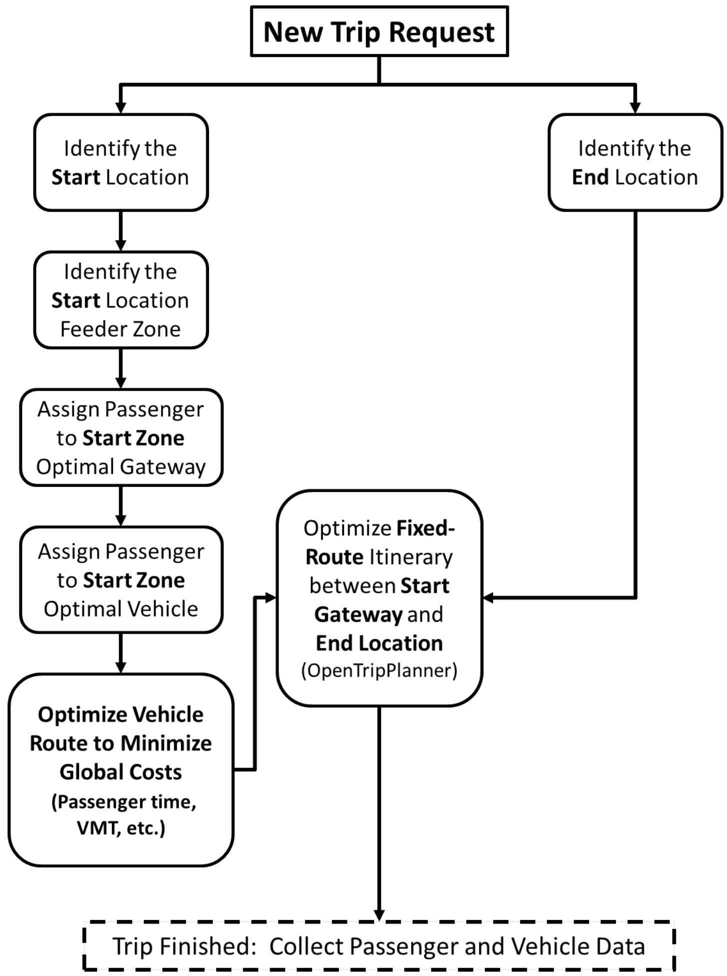

Figure 4-4: Data flow of a trip that uses demand-responsive transportation for the first leg and fixed-route for the rest of the trip.

4.2.3.3 Fixed-Rotue to Demand-Responsive Trips

In fixed-route to demand-responsive (FD) trips, the first leg of the trip is handled by the fixed-route transit network and the final leg or "last mile" of the trip is handled using demand-responsive transportation. The data flow within the NITS simulator for an FD trip is shown in Figure 4-5. Once a destination gateway is chosen, an optimal fixed-route trip is found using OpenTripPlanner. Upon arriving at the destination gateway, the passenger's optimal demand-responsive vehicle and route are selected.

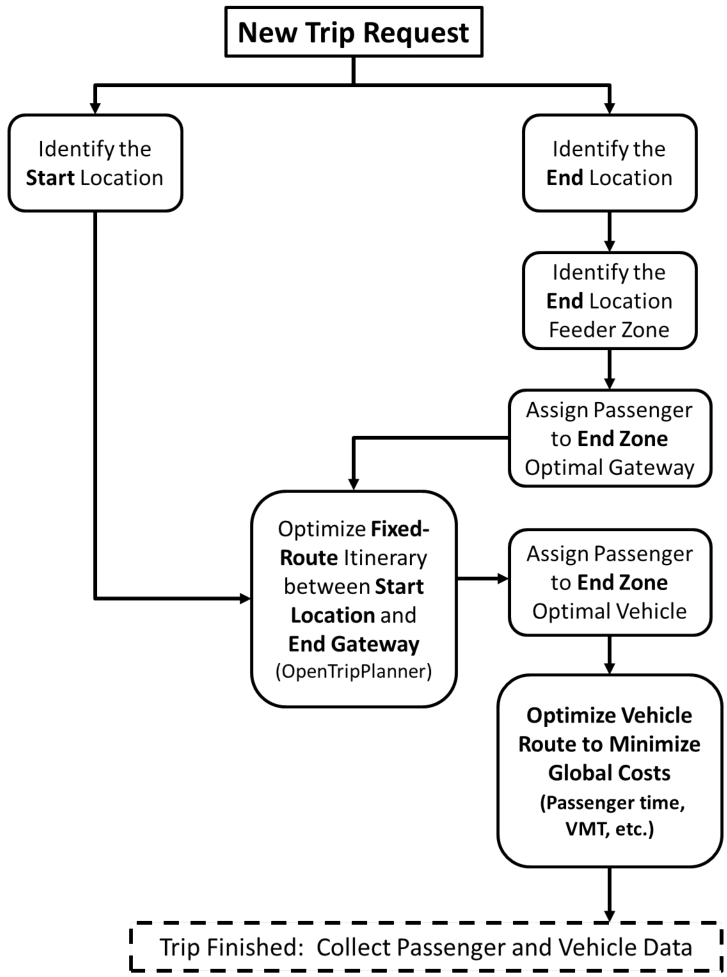

Figure 4-5: Data flow of a trip that uses fixed-route transportation for the first leg and demand-responsive transportation for the final leg.

4.3 Operation of the NITS Simulator Components

The flowcharts shown in Section 4.2 show the basic flow of data through the NITS simulator for each type of trip supported by the simulator. Each of the boxes in in these figures represents a decision or optimization task. Many of these optimization tasks are complex and justify further explanation. This section breaks down the logic and algorithms of the individual NITS simulator components.

4.3.1 Fixed-Route Itinerary Optimizer (OpenTripPlanner)

In order to create optimal passenger itineraries consisting of both fixed-route and demand-responsive transit trips, a method of optimizing the fixed-route portion of the trip is required. Fortunately, many such optimizers exist and are freely available for use. Fixed-route optimizers considered for this research include Google Transit, Ovi Maps, Hop Stop, and OpenTripPlanner. OpenTripPlanner (OTP) was selected for this research due to its ability to be licensed as an open-source project and its ability to find optimal routes in real-time [63].

OpenTripPlanner accepts as input a passenger's start location, ending location, and starting time. OTP returns the optimal itinerary for that passenger between those two points and the time required to complete the journey. For the purpose of this research, optimal refers to minimum travel time. The travel times are found using an A* search algorithm that parses a general transit feed specification to find optimal transit routes and searches the Open Street Map graph in order to include information on walking time and distance [51].

This information is used for two purposes. The first purpose is to provide a baseline for travel time. Using the existing fixed-route infrastructure in a city, OpenTripPlanner provides the travel time needed by each passenger. This travel time acts as the point of comparison to any improvements made by introducing demand-responsive transportation.

The second use of OTP is to provide information on travel time for individual legs of a passenger's journey that require fixed-route transportation. For instance, if the first and final legs of a passenger's journey uses demand-responsive transportation, the middle leg of the journey will be handled with fixed-route transportation, most likely rail or BRT. Information on this middle leg will be provided by OpenTripPlanner.

4.3.2 Point-to-Point Vehicle Optimizer

In addition to obtaining optimal fixed-route transit routing, the simulator also requires optimal point-to-point vehicle routing. Point-to-point routing refers to finding the optimal bus route between two geographic points. Many such point-to-point routers exist. This research utilizes an open-source routing engine called the Open Source Routing Machine (OSRM) [47]. The OSRM was developed by Sanders et al. [43] at Karlsruher Institute of Technology to solve continent-wide routing problems very quickly by utilizing hierarchical highway node and transit node routing algorithms as well as efficient contraction hierarchies [45] [44].

The research conducted for this thesis utilizes the OSRM to identify optimal vehicle routes for a demand-responsive fleet of vehicles. Given two points on a map, the OSRM will quickly identify the fastest route between these two points. This information is utilized by the dial-a-ride vehicle optimizer to determine which order the passenger starting and ending locations should be visited.

4.3.3 Demand-Responsive Vehicle Optimizer

Solving the point-to-point vehicle routing problem represents only a small portion of the dial-a-ride problem (DARP). While the vehicle routing problems seeks to find the optimal route between two points, the objective of the DARP is the find the optimal order in which to visit the passenger locations. For example, given five passengers, there are ten locations to visit, a pickup and drop-off point for each passenger. The point-to-point optimizer provides the shortest path between any two of the ten points while the DARP solver finds the optimal order in which to visit those ten points.

The simulator used in this research uses a custom-built, python-based, dial-a-ride problem solver. The dial-a-ride problem being solved in this research is the online dial-a-ride problem discussed in Section 2.1.2, meaning that the location and request times of passengers are not known *a priori*. As each new passenger makes a request, that passenger is assigned to a vehicle, and that vehicle must update its route in order to accommodate the passenger's request while minimizing global costs.

One question that may arise is to wonder why the simulator is solving the online instead of the offline dial-a-ride problem. A simulator knows all the passenger requests *a priori*, which means that the dial-a-ride solver can use this information to find better routes. The reason that this is not done is because it does not reflect a practical implementation. The simulator is intended to simulate the behavior of passengers and vehicles in a real-world application. In real-world applications, *a priori* information is not known. Using *a priori* information in a simulator would lead to results that cannot be reproduced in a practical application.

Two approaches to solving the dial-a-ride problem were utilized in the research. The first approach is a genetic algorithm (GA) approach. In the GA approach, each time a passenger is added to a vehicle, a new order for visiting each passenger location is found. An example solution to the DARP found by the GA approach is shown in Figure 4-6. In the example solution, the order in which five passenger locations are visited is shown. The number represents the passenger number, and the 'P' and 'D' notation indicates whether this stop is picking up or dropping off a passenger. Reading from left to right, the solution in Figure 4-6 shows that the vehicle will pick up passenger one first, pickup passenger three second, then drop off passenger three, and so on until passenger five is dropped off last.

$$[1P, 3P, 3D, 5P, 2P, 4P, 2D, 4D, 1D, 5D]$$

Figure 4-6: Sample solution of the dial-a-ride problem.

The GA solver will iterate through several generations of similar solutions until a near-optimal solution is found. This approach is capable of finding near-optimal solutions, but the drawback to this approach is that the algorithm is very expensive in terms of computation time. In a large system with dozens of vehicles and hundreds of passengers, the genetic algorithm may spend several minutes search for possible routes. Since this task must be repeated each time a passenger enters the system, this time cost is not feasible to implement in a real-world scenario.

The second approach is a heuristic method based on an exhaustive-search insertion method outlined by Horn in [30]. The insertion method operates much more quickly than the GA approach. The insertion method reduces the search space by locking the order in which passenger locations are visited once they are set. This means that each time a passenger is inserted and an optimal solution is found, that order cannot be changed when future passengers are inserted. When a passenger requests a trip, that passenger's starting and ending destination will be inserted between the stops that were ordered previously. For instance, if a new passenger requests a trip for the vehicle serving the passengers shown in Figure 4-6, then the possible insertion points for the new passenger are show in in Figure 4-7. Each space in Figure 4-7 represents a possible insertion point for the new passenger's pickup and drop-off location.

$$[\_\_, 1P, \_\_, 3P, \_\_, 3D, \_\_, 5P, \_\_, 2P, \_\_, 4P, \_\_, 2D, \_\_, 4D, \_\_, 1D, \_\_, 5D, \_\_]$$

Figure 4-7: Possible insertion points for the next passenger's origin and destination.

Given that there are only 2N+1 possible insertion points for the pickup and drop-off locations of the new passenger, the exhaustive-search insertion method will iterate over every possible insertion point and select the optimal solution.

The optimal solution for both of these optimization approaches is shown in Equation 4.1. The definitions of each of these variables are provided in Table 4-1.

$$\underset{P}{\operatorname{argmin}}\, d_j(P) = \underset{P}{\operatorname{argmin}} \left(\alpha VMT_{j,P} + \beta \sum_{i=1}^{N} (w_{wt}T_{wt,i,P} + w_{rd}T_{rd,i,P}) \right) \tag{4.1}$$

Table 4-1: Variable definitions for the demand-responsive optimization problem.

Variable	Definition
P	a possible solution to the DARP, such as that shown in Figure 4-6
$d_j(P)$	the global cost of operating vehicle *j* given solution *P*
$VMT_{j,P}$	the vehicle miles traveled of vehicle *j* given solution
α	a scalar set by the operator on the importance of VMT cost
$T_{wt,i,P}$	the amount of time that passenger *i* will spend waiting for a vehicle given solution *P*.
w_{wt}	a scalar representing the weight of passenger wait time
$T_{rd,i,P}$	the amount of time that passenger *i* will spend riding a vehicle given solution P
w_{rd}	a scalar representing the weight of passenger ride time
β	a scalar set by the operator on the importance of passenger convenience costs

As each new passenger is inserted into the simulation, a combination of operator costs and passenger costs is minimized. In Equation 4.1, *P* represents a potential path that a vehicle can take to visit a set of passenger locations. The object of the DARP solver is to identify the P that minimizes the VMT of the vehicle as well as the passenger costs of each passenger being serviced by that vehicle.

4.3.4 Optimal Passenger Vehicle Assignment

The algorithm to determine the optimal demand-responsive vehicle to service a passenger requests relies upon the dial-a-ride problem solver discussed in the previous section. One advantage of using the NITS framework is that the individual demand-responsive zones are kept relatively small and relatively few vehicles are utilized within each zone. For instance, in city-wide simulations for the city of Atlanta, feeder zones rarely utilized more than 10 vehicles. Because of this, an exhaustive search algorithm is used to determine which vehicle within a given feeder zone is the optimal vehicle to service a new passenger request.

Each time a passenger requests a trip with in a zone, each vehicle is considered as a potential vehicle to service that request. The passenger is temporarily assigned to each vehicle in the coverage area. Then the DARP solver is run for each vehicle to determine the global costs incurred if that trip were actually assigned to that vehicle. The vehicle with the lowest global cost is selected as the optimal vehicle for that trip. Equation 4.2 describes this mathematically, where $\underset{P}{\operatorname{argmin}}\, d_j(P)$ is the minimum path found using Equation 4.1 and $d_{\phi_k}(j)$ is the cost of utilizing vehicle j within subnet ϕ_k to handle the new trip.

$$\underset{d \in \phi_k}{\operatorname{argmin}}\, d_{\phi_k}(j) = \underset{d \in \phi_k}{\operatorname{argmin}} \left(\underset{P}{\operatorname{argmin}}\, d_j(P) \right) \tag{4.2}$$

4.3.5 Optimal Passenger Gateway Assignment

Similar to selecting an optimal vehicle, selecting an optimal gateway builds upon lower levels of optimizations. In certain setups, a feeder area can have more than one possible gateway. In these scenarios, it is important to identify which gateway would be the most efficient. Once again this, the algorithm for choosing the optimal gateway takes advantage of the fast DARP solver to perform an exhaustive search.

For every possible gateway, the optimal vehicle and route are calculated. The gateway with the lowest global cost is selected. Equation 4.3 illustrates this mathematically. Notice that $\underset{d \in \phi_k}{\operatorname{argmin}}\, d_{\phi_k}(j)$ is the minimum cost vehicle selected using

Equation 4.2 and $d_{\gamma_{\phi,k}}(l)$ is the cost of selecting gateway l within the set of possible gateways $\gamma_{\phi,k}$.

$$\underset{l \in \gamma_{\phi,k}}{\operatorname{argmin}}\, d_{\gamma_{\phi,k}}(l) = \underset{l \in \gamma_{\phi,k}}{\operatorname{argmin}} \left(\underset{d \in \phi_k}{\operatorname{argmin}}\, d_{\phi_k}(j) \right) \tag{4.3}$$

In a very large network with many vehicles and many possible gateways, this approach may not be computationally efficient. However, given the motivation for this research is the city of Atlanta which contains only 38 possible gateways, this is the total number of rail stations in the MARTA system, and a limited number of vehicles required to operate within each feeder zone, an exhaustive search approach is not infeasible given the availability of a fast dial-a-ride problem solver. For M possible gateways within a zone and N possible vehicles, there is a total of MN possible options. Given that a typical setup M<10 and N<10, meaning that MN<100, an exhaustive search is feasible.

4.3.6 Identifying the Demand-Responsive Zones

Identifying the demand-responsive zone for the passenger locations is handled using geo-fencing. Each demand-responsive zone is defined by a set of ordered latitude and longitude pairs forming a polygon around the service area. To determine which demand-responsive feeder zone a passenger location belongs to, each feeder zone is independently tested to see if the passenger lies within the zone. If a passenger location is determined to be within the geofence of a zone, the passenger is assigned to that zone. If a passenger is within more than one geo-fence, the union of each of those geo-fences is treated like a super zone for that passenger. The effect of this is simply that more possible gateways and vehicles will be considered for each passenger. In most scenarios this will not cause problems with performance because the number of gateways and demand-responsive zones are low. For instance, in the simulations conducted for this research, the total number of rail stations acting as gateways is 38. This is every rail station in the city of Atlanta, which is a low enough number that the algorithm can exhaustively check each feeder zone and gateway for optimality.

4.4 Handling Disruptions to Service

The simulator built for this research is designed to be easily modified into a practical integrated demand-responsive dispatching agent. For this reason, it cannot be assumed that all vehicles will operate at 100% efficiency at all times. In real-world conditions, malfunctions and vehicle breakdowns are bound to occur. The simulator must be able handle these situations and gracefully recover from them. In the event of a vehicle failure, all passengers assigned to that vehicle will need to be rerouted. Each passenger will be reinserted into the system with a new trip request and new parameters for location and time. While the ending location of each passenger will not be altered, if a passenger has already been picked up by the vehicle, that passenger's new starting location will be the nearest safe waiting spot to the malfunctioning vehicle. The passenger's cost function will also be changed. Any waiting, riding or walking time accumulated by the passenger at the point of the vehicle breakdown is added to the passenger's total trip cost. This ensures that these passengers will be given priority over new passenger's entering the system. Using a non-linear cost function that places increasing penalties for longer wait times can emphasize this priority. The routing, vehicle assignment, and vehicle dispatching algorithms will run normally. In the event that these new passenger trips require a new vehicle, then the vehicle will be dispatched according to the normal dispatching algorithm that is run each time a new passenger enters the system.

A second type of disruption, and a type that is far more common, is that of schedule deviations in fixed-route vehicles. The simulator uses the OpenTripPlanner to calculate the optimal fixed-route itinerary for each passenger. In the event that a bus or train is late or early, these itineraries will change. The simulator described in this chapter is designed to accept updated changes from OpenTripPlanner if vehicles deviate from the schedule. However, real-time updates to the OpenTripPlanner optimizer are still under development in the OpenTripPlanner community and not available for practical use. Since current work focuses on offline simulations, this is not currently a concern. However, in a practical application schedule times cannot be assumed and real-time updates need to be considered.

One final set of disruptions that must be considered is that of traffic bottlenecks, road closures, or accidents that affect the street network. The point-to-point vehicle router relies on the Open Source Routing Machine (OSRM) for optimal route selection. The current implementation of the OSRM assumes an average speed for each road that does not take into account real time traffic data. While a multiplier for travel times can be used as an approximation for travel at various times of the day, real-time information such as road closures or accidents are not considered. These types of disruptions can have significant effects on demand-responsive vehicle routes. In fact, one major advantage of demand-responsive vehicles versus fixed-route vehicles is that they can adjust to these types of disruptions in real time. As with the OpenTripPlanner, integrating real-time traffic information with Open Street Map is a current topic of interest for many developers and researchers [64]. While the simulation outlined in this chapter does not currently utilize real-time information, it is poised to accept this data as it becomes available.

4.5 Information Required by the NITS Simulator

The previous sections outline the basic data flow within the NITS simulator and provide an analysis of the operation of each component within the simulator. This section will describe the configuration information and data required by the simulator in order to simulate a demand-responsive service. This information includes the following.

1. A set of demand-responsive zones are defined by the system operator. These zones are defined by a virtual perimeter, or geo-fence, consisting of a set of latitude and longitude pairs. The area within the geofence is considered to be a demand-responsive service area while the area outside the geo-fence operates only fixed routes.
2. For each geofence, one or more transit gateways must be defined. A transit gateway is an interface between the demand-responsive vehicles and the large fixed-route transit network. If a passenger wishes to leave the demand-responsive area, he or she is taken to one of the gateways to be transferred to the fixed-route network. Typically these gateways are rail or bus rapid transit stations.
3. A general transit feed specification (GTFS) file must be provided to the NITS simulator. The GTFS file contains all of the fixed-route information and schedules available to the passenger. This information is used by the transit route optimization routine to find optimal fixed-route itineraries for each passenger. In a typical NITS setup, nearly every passenger trip will utilize fixed-route transit in some way. Some passengers may use the fixed-route network for the entirety of their trip, while others may use it in conjunction with demand-responsive transit at the first or last miles of the passenger's trip. GTFS data for a large number of transit agencies can be found at gtfs-data-exchange.com [49].
4. A street map must be uploaded into the simulator. The NITS simulator requires information about the street network in order to optimally route demand-responsive vehicles between any two points on a map as well as to

solve the dial-a-ride problem for selecting which order in which to visit passengers. Street map data can be obtained by downloading an .osm file from Open Street Map [38].

5. A set of passenger data is needed for simulation. The passenger data must include a starting location, ending location, and desired trip start time for each passenger. This research utilizes a set of passenger survey data provided by the Atlanta Regional Commission to provide an accurate representation of transit demand in the city of Atlanta. If direct passenger data is not available, passenger trips can be simulated using travel demand modeling.

Given these 5 inputs, the NITS simulator can simulate the exact itinerary for every passenger and return to the transportation engineer the cost and time associated with each trip.

4.6 Verification of Simulation Results

As described in the previous sections of this chapter, some constituent components of the NITS simulator were developed by 3rd party developers to provide point-to-point and fixed-route optimization while custom-written extensions were created to generate integrated transportation itineraries. Since this research relies solely on the results generated by this complex transit simulator, it is imperative that the simulator provides accurate and realistic results. The final section of this chapter validates the fixed-route and demand-responsive vehicle results generated by the NITS simulator

4.6.1 Verification of Vehicle Route Results

The vehicle routes selected by the optimizer should be feasible. Feasibility, in this case, means that the vehicle routes and stop visit times that are selected by the optimizer can be completed within the stated amount of time. If the simulator and optimizer are assigning routes that are not realistic, then the results delivered by the simulator are not useful for designing real demand-responsive systems.

To test for feasibility, the itineraries that were selected for a set of demand-responsive vehicles are passed through three 3rd party route optimizers, Google Maps, Bing Maps, and MapQuest. Each of the itineraries contains between 5 and 15 passenger stops dispersed throughout Midtown Atlanta. The order of the stops was selected by the NITS optimizer. For this set of bus itineraries, the NITS planner calculated that the average time for a vehicle to complete each itinerary is 691 seconds. Without changing the order of any of the itineraries, they were all passed through the three 3rd party route optimizers. Google Maps calculated the average completion time to be 680 seconds. Bing Maps estimated the travel times to be 605 seconds, and MapQuest estimated the travel times to be 405 seconds. These results are compared in Figure 4-8.

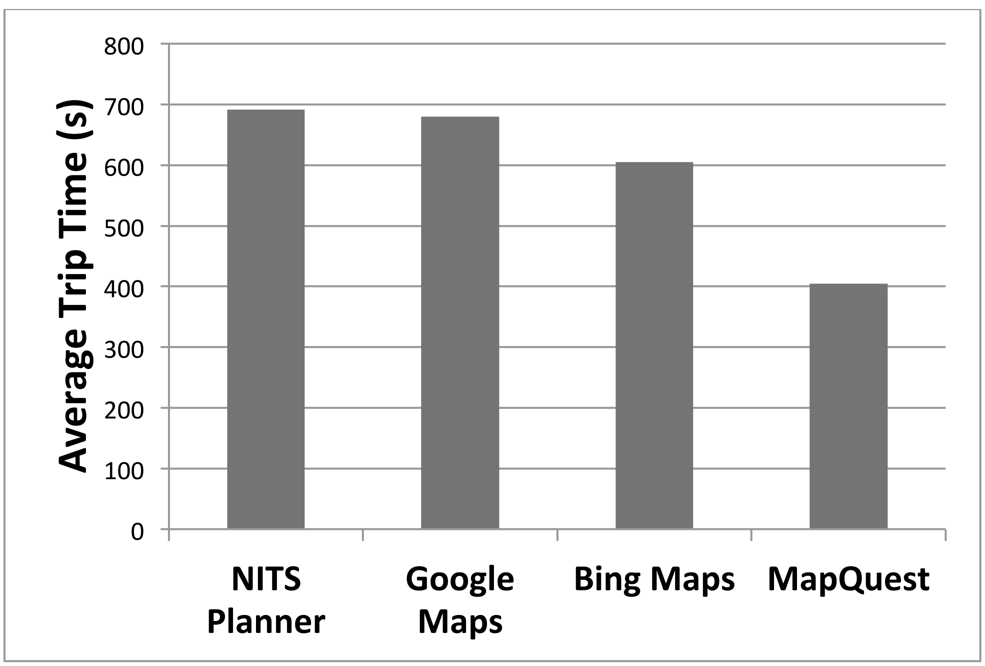

Figure 4-8: Comparison of average route times for four route optimizers.

These results show that the route times estimated by the NITS planner using the Open Source Routing Machine are in line with other widely accepted route planning tools. The results of the NITS planner show slightly longer travel times because the results delivered from the NITS optimizer are scaled upward slightly to account for passenger loading and unloading as well as the fact that a van or small bus is less maneuverable in traffic than a standard car.

4.6.2 Verification of Passenger Itineraries

In addition to the vehicle itineraries, passenger itineraries must also be feasible and realistic. Passenger itineraries can consist of demand-responsive trips as well as fixed-route or walking trips. Unlike the vehicle itineraries, it is not possible to directly compare passenger travel times in a demand-responsive trip against a 3rd party optimizer because the times that each passenger spends in the vehicle is largely dependent upon other passengers in the simulation as well as the optimization algorithm. However, some checks can be applied to compare the time that each passenger spent inside of a demand-responsive vehicle is at least as long as the time that it would take the passenger to drive directly between his or her origin and destination. For one simulation, a set of passenger ride times were compared with driving times between each passenger's origin and destination. The time that each passenger spent inside of a demand-responsive vehicle averaged 41% longer than the time that it would take each passenger to drive that same distance directly. There were zero demand-responsive trips that took less time than directly driving those same trips. If any such trips were found, it would indicate an error in the model.

The transit portion of each trip was also examined for reliability. For a simulation in Midtown, Atlanta, the recorded times to complete the necessary transit trips were compared with results given by a 3rd party optimizer. The 3rd party optimizer in this case was Google Transit. The transit times calculated by the OpenTripPlanner used in the

NITS simulator were consistently within 3%, on average, of the times calculated by the Google Transit optimizer.

The passenger itineraries, demand-responsive vehicle routes, and fixed-route travel times are all consistent with existing, and widely-accepted, third-party optimizers. This fact demonstrates that the results delivered by the NITS simulator are accurate and useful for providing the necessary passenger and operation data to optimize and compare transit performance in real-world scenarios.

It is imperative that the simulator described in this chapter provides accurate results for a wide range of settings and implementations. The simulator is used as the testbed for the remainder of the research outlined in this thesis. In the next chapter, the simulator will be used to compare passenger satisfaction and operator costs in fixed-route and demand-responsive transportation settings in order to determine when and where demand-responsive transportation can be effectively implemented. In Chapter 6, the simulator will be used to find the optimal size and shape of demand-responsive feeder systems by testing a wide array of feeder system designs to determine which design is optimal for a given implementation.

CHAPTER 5

COMPARING FIXED-ROUTE AND DEMAND-RESPONSIVE TRANSPORTATION PERFORMANCE

Since cities can vary significantly in density from one area to another, it is important that transportation planners are able to determine which parts of a city are better suited to demand-responsive transportation and which parts are better suited to fixed-route transportation. To make this determination, an objective method of comparing the performance of demand-responsive transportation and fixed-route transportation is needed. In this chapter, a generic method of objectively comparing demand-responsive and fixed-route transportation performance will be introduced. The method will adapt the comparison techniques that were developed by Quadrifoglio and others, which were discussed in Section 2.2, and will extend those techniques to heterogeneous environments in which passenger behavior, road networks, and transit schedules are difficult to model. The comparison technique introduced in this chapter will take advantage of the NITS framework that was outlined in Chapter 3 and will utilize the demand-responsive transportation simulator that was introduced in Chapter 4.

This chapter is divided into three sections. In section 5.1, a case study is presented that demonstrates how fixed-route and demand-responsive transportation performance is compared in a homogeneous city layout. The purpose of this section is to describe existing comparision techniques in greater detail than is presented in the literature review and to demonstrate the type of results that the comparison techniques should produce. In section 5.2, a novel, simulation-based comparison technique is presented that will provide accurate comparisions in any type of city layout. In section 5.3, the simulation-based comparison technique is used to compare the performance of a proposed demand-responsive feeder system for the City of Atlanta with the performance of the existing fixed-route system. The case-study will use actual passenger data, Metro Atlanta Rapid Transit Authority (MARTA) schedules, and open-source street maps for the City of Atlanta to provide accurate results on how a demand-responsive system would perform.

5.1 Comparison of DRT and FRT in a Gridded Street Layout

In the first of two case studies, fixed-route and demand-responsive transportation performance is analyzed in a homogoeneous city. This case study provides a step-by-step walk-through of how demand-responsive and fixed-route transportation performance is analyzed and illustrates how these comparisons can be used to influence transportation policy.

In this case-study, three different fixed-route systems and one demand-responsive system are compared in a fictional city with a gridded street layout, regular transit schedule, and uniform passenger dispersion. The passenger costs and operator costs for each system are derived and the combined passenger and operator costs are compared for various levels of passenger demand. As passenger demand increases, the cost of operating demand-responsive transportation also increases. At a certain point, the passenger demand level becomes too expensive to operate a demand-responsive system and fixed-route service becomes the more optimal solution. In this case study, the four different systems are compared and the point at which fixed-route transit becomes more efficient than demand-responsive transit is determined.

5.1.1 Layout and Behavior of the Gridded Street System

In the fictional city, the street layout follows a perfect grid, a set of buses operates at unchanging intervals within the grid, and passengers are uniformly dispersed throughout the city. The bus layout of the city under examination is a grid consisting of NxN bus stops evenly separated by a distance H, as seen in Figure 5-1. The grid of bus stops are serviced by a set of N north/south bus lines and a set of N east/west bus lines traveling at a speed of v_f with headway τ.

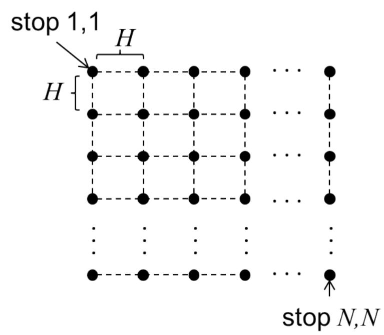

Figure 5-1: Bus stop layout in gridded street simulation.

In the fixed-route system, passengers walk to the nearest bus stop, wait on the bus, possibly transfer between buses, ride a second bus, and then walk to their final destinations. In the demand-responsive system, a demand-responsive vehicle carries the passenger from his or her origin to the nearest fixed-route bus stop, the passenger will use fixed-route transit for the middle leg of the journey, and demand-responsive vehicle will handle transporting the passenger between the passenger's final bus stop and his or her final destination. In this scenario, the set of demand-responsive vehicles act as a feeder system for the fixed-route bus network.

The purpose of this examination is to determine whether a demand-responsive feeder system with sparse fixed-route bus stops is preferable to a purely fixed-route system with denser bus stops within the system. As with Li and Quadrifoglio [37], a combined cost of passenger walk time, wait time, and transit time will be considered as the passenger costs, and vehicles miles traveled (VMT) will represent operator costs.

5.1.2 Fixed-Route Passenger Costs

To create an efficient transit system, passenger costs and operator costs must be minimized. Since this case study is dealing with a homogoneous street layout, it is possible to find closed-form solutions to represent passenger and operator costs. In this example, passenger costs consists of walking time T_{wk}, waiting time T_{wt}, and transit time T_{tr}. The combined passenger costs, J_P, is represented as a weighted combination of these individual costs.

$$J_P = w_{wk} \times T_{wk} + w_{wt} \times T_{wt} + w_{rd} \times T_{tr} \tag{5.1}$$

5.1.2.1 Passenger Walking Time

For the fixed-route system, passenger locations are spread uniformly across the city, and will always walk to the bus station closest to their origin location. Therefore, the expected distance that a passenger will walk, $E(D_{wk})$, is a combination of that passenger's north/south walking distance, $E(D_{wk,v})$, and his or her east/west walking distance, $E(D_{wk,h})$.

$$E(D_{wk}) = E(D_{wk,h}) + E(D_{wk,v}) \tag{5.2}$$

With square coverage areas, the average time spent walking in a vertical direction is equal the average time spent walking in a horizontal directin, $E(D_{wk,h}) = E(D_{wk,v})$, which means that $E(D_{wk}) = 2E(D_{wk,h})$.

Since the station is located in the center of the feeder zone, the range of distances that passengers can walk is [0,*H*/2], and since the passengers are uniformly spaced along this interval, $E(D_{wk,h}) = H/4$. Therefore combining the expected vertical and horizontal waking distances for the passenger yields,

$$E(D_{wk}) = 2E(D_{wk,h}) = H/2. \tag{5.3}$$

This represents the total expected walking distance between the passenger's origin and the nearest bus stop. To find the total expected walking distance for the trip, this number must be doubled since a passenger must also walk between his or her final bus stop and to the destination.

To find the total expected passenger walking time, divide the expected distance by the average passenger walking velocity.

$$E(T_{wk}) = E(D_{wk})\frac{1}{v_w} \tag{5.4}$$

5.1.2.2 Passenger Waiting Time

Since passengers request trips at a uniform rate, the range of wait times that a passenger must endure is $[0,\tau]$, where τ is the average time between subsequent buses, i.e., bus headway. With a uniform arrival time, the expected wait time for any bus is, $E(T_{wt}) = \tau/2$. In the grid system, most trips will result in a rail transfer causing the passenger to experience two waiting periods. In a system with *N* vertical and *N* horizontal bus lines, the probability that a passenger will require a transfer is (*N*-1)/*N*. Therefore the total expected wait time for a trip is,

$$E(T_{wt}) = \frac{\tau}{2} + \frac{(N-1)\tau}{2N} = \frac{(2N-1)\tau}{2N}. \tag{5.5}$$

The value, $\frac{(N-1)\tau}{2N}$, represents the probabilty of requiring a transfer multiplied by the expected wait time for that transfer. The value, $\frac{\tau}{2}$, represents the expected wait time for the first bus.

5.1.2.3 Passenger Transit Time

Similar to passenger walking time, passenger ride time is composed of a vertical and horizontal component,

$$E(D_{tr}) = E(D_{tr,h}) + E(D_{tr,v}). \tag{5.6}$$

For a symmetric, *N*x*N*, system the expected ride time is simplified to $E(D_{tr}) = 2E(D_{tr,h})$. To find $E(D_{tr,h})$, consider a single horizontal bus line with *N* stops. Along this horizontal bus line, there are *N* possible starting locations and *N* possible destination locations. For each possible starting location *i,* there is an expected transit time $E(D_{tr,h,i})$. Assuming that the destination bus stop for every passenger is uniformly dispersed, there is an equal chance that a passenger's destination will be any of the other stops along the line, and the expected distance is found by,

$$E(D_{tr,h,i}) = \frac{1}{N}\sum_{j=1}^{N} |i-j|H, \tag{5.7}$$

for,

$$i \in 1, N.$$

From equation 5.7, the expected horizontal travel distance without specifying the start location is found by,

$$E(D_{tr,h}) = \frac{1}{N}\sum_{i=1}^{N} E(D_{tr,h,i}) = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} |i-j|H. \tag{5.8}$$

The single line transit time is then found by dividing the expected distance by the average bus speed.

$$E(T_{tr,h}) = E(D_{tr,h})\frac{1}{v_f} \tag{5.9}$$

The total transit time, vertical and horizontal, is found by doubling the expected single line ride time.

$$E(T_{tr}) = 2E(T_{tr,h}) \tag{5.10}$$

5.1.3 Fixed-Route Operator Costs

The cost incurred by the operator, J_O, is represented as the total vehicle miles traveled (VMT) of the fleet, which is a combination of the fixed-route VMT and dynamic-route VMT.

$$J_O = VMT_f + VMT_d \tag{5.11}$$

Fixed route VMT is found by summing the travel distance of each bus, on every line across the entire simulation time.

$$VMT_f = 2(N+N)\frac{T}{\tau}(NH) \tag{5.12}$$

Where, *(N+N)* is the total number of bus lines, $\frac{T}{\tau}$ is the length of the simulation divided by the average headway, and *NH* is the length of the bus line. The product of these values is then doubled to account for buses traveling in both directions, (i.e., the north/south lines has buses traveling in both the northbound and southbound directions and the east/west lines have buses travling in both the eastbound and westbound directions). Equation 11 can then be reduced to,

$$VMT_f = 4N^2 \frac{T}{\tau} H. \tag{5.13}$$

5.1.4 Demand-Responsive Transportation Costs

The closed-form solutions above represent the expected passenger wait, walk, and transit time for all fixed-route passengers within the system as well as the expected vehicle miles traveled by fixed-route route vehicles. These closed-from formulas allow for quick analysis of various fixed-route solutions. If the distance between bus stops is increased or decreased or if the average headways are made shorter or longer, then these new characteristics can be easily updated in the formulas and updated passenger costs and operator costs can be found.

For the demand-responsive option, closed-form solutions are less likely to provide accurate measures of passenger costs or expected mileage covered by the demand-responsive vehicle. Finding a relationship between passenger wait times and ridership demand is a non-trivial task and this relationship does not necessarily hold when the environment is updated in any manner. Therefore, this study will utilize the simulator described in Chapter 4 to simulate the actual routes of vehicles as well as the actual passenger wait and ride time. A gridded street network is loaded into the simulator and passengers are created according to the specified passenger demand rate and a uniform distribution of passenger locations and passenger request times is assumed. The passenger wait, walk, and ride times as well as the VMT of the demand-responsive and fixed-route vehicles are recorded by the simulator and are used to determine the total cost of operating demand-responsive transportation.

5.1.5 Simulating Demand-Responsive Costs

The cost functions derived in the previous section are generic formulas for homogoneous gridded street system. In order to perform analysis, specific characteristics of the city must be provided. In this example, the characteristics of the system are:

- N = number of horizontal and vertical transit lines = 6,
- H= block length = 4400 ft (1342 m),
- v_w= walking velocity = 3 mph (4.83 kph) = 4.4 ft/s (1.342 m/s),
- v_f= average bus velocity = 15 mph (24.15 kph) = 22 ft/s (6.71 m/s),
- v_d= average dynamic-route vehicle velocity = 10 mph (16.10 kph) = 14.7 ft/s (4.48 m/s), and
- τ= average headway = 600s.

The transportation engineer for this city must analyze and compare four options for improving transit performance. The four choices are as follows.

Scenario 1: Use a DRT system where a single demand-responsive vehicle is utilized at each bus station to act as a feeder for the station.

Scenario 2: Decrease the distance between each bus station to 2200 ft (*H*=2200) to make the distance more walkable.

Scenario 3: Decrease the distance between each bus station to 3300 ft (*H*=3300).

Scenario 4: Make no change to the system, leaving the distance between each station at 4400 ft (*H*=4400).

In these scenarios, the passenger cost is defined as the average door-to-door travel time, and the cost to the system operator is defined as the total vehicles miles traveled

(VMT) of the fleet. The total cost of the system is a scaled average of the passenger cost and the operator cost.

For Scenarios 2, 3, and 4, the fixed-route costs were calculated using the formulas dervied in Sections 5.1.2 and 5.1.3. The demand-responsive costs from Scenario 1 were caluclated using the demand-responsive transportation simulator. Since demand-responsive performance is highly correlated with passenger demand, multiple simulations were run for various levels of passenger demand.

The results of this analysis and simulations are shown in Figures 5-2 through 5-5. In Figure 5-2, the total average door-to-door trip time of the passengers is shown for five different levels of passenger demand. The passenger demand ranges from 1.5 passengers per minute to 15 passengers per minute. Since fixed-route transportation schedules are independent of real-time demand, the costs of operating the three fixed-route scenarios are shown as flat lines across the various passenger demand levels. The curved line shows the average travel time for passengers in Scenario 1. This line is not flat because demand-responsive transportation performance is dependent upon the passenger demand. As more passengers enter the system wait times incraese for all passegners and the total VMT of the fleet increases. This graph shows that for very low levels of passenger demand, a DRT system outperforms all the fixed-route systems. However, as passenger demand increases, fixed route service with frequent stops becomes more effective. The point at which the average door-to-door travel time of Scenario 2 exceeds that of Scenario 1 is at approximately 2 passengers per minute entering the system. The point at which travel time in Scenario 3 exceeds that of the DRT system is at approximately 9 passengers per minute. For these simulations, Scenario 4 failed to improve travel time over the DRT system for any level of passenger demand.

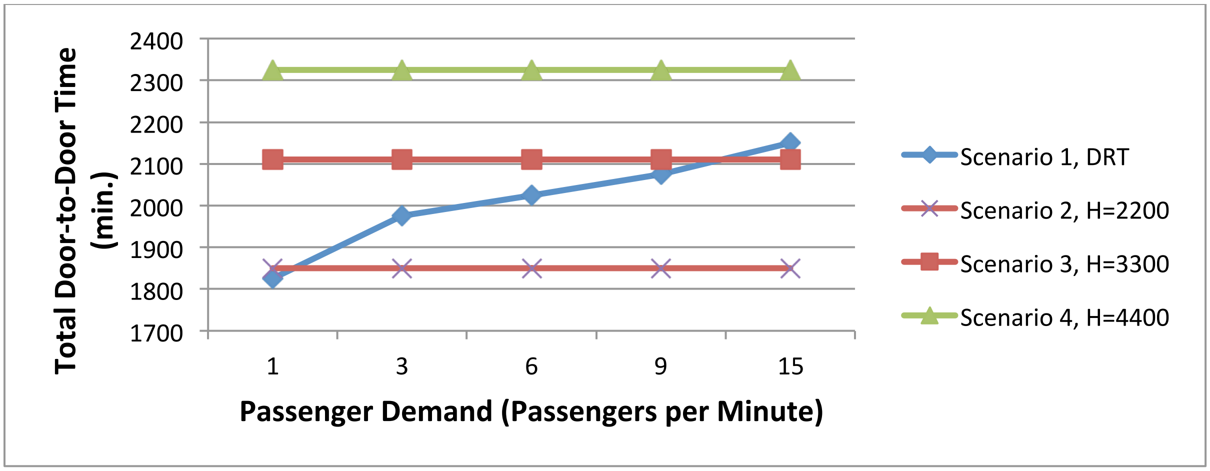

Figure 5-2: Total Door-to-Door Travel Time

Using passenger door-to-door travel time as the sole metric for selecting a transportation system would indicate that Scenario 2 is the optimal solution for enarly all passenger demand levels. However, decreasing door-to-door travel time comes at a cost to the system operator. Since operator costs cannot be ignored, operator costs must be averaged in with the passenger travel time to provide a more realistic representation of total system cost. Figure 5-3 shows the travel time costs for the passenger combined with the operation cost (VMT). The total VMT of the fleet, as well as the average passenger time for each scnenario, are scaled to values between 0 and 1 so that they can be compared. The total scaled cost of passenger wait times and VMT will be a value between 0 and 2. Notice Scenario 3 (H=3300). While this scenario was not the best choice for travel time, when the cost of VMT is factored in, it becomes the most effective of the three fixed-route scenarios.

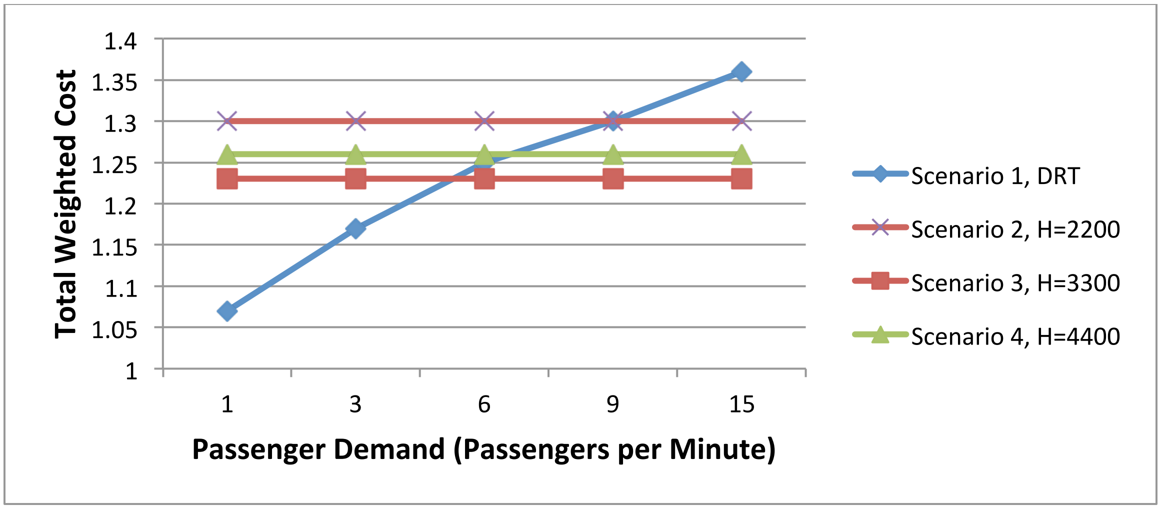

Figure 5-3: Combined VMT and Passenger Costs in the System.

When considering combined passenger and operator costs, demand-responsive transportation is more effective than any of the fixed-route options for any passener demand level below 6 passengers per minute. At rates above six passengers per minute, the best policy is Scenario 3. Depending on the time of day and the passenger demand levels at different portions of the city, Scenario 1 or Scenario 3 should be selected for the transportation system.

This example showed how demand-responsive and fixed-route transportation can be objectively compared across a city-wide feeder system with a homogenous layout. Fixed-route passenger and operator costs were represented by closed-form solutions and demand-responsive costs were calculated through simulation. The example is simple, but illustrates he basics of how demand-responsive and fixed-route comparisons can be made. However, since the goal of this research is to examine the possibility of using demand-responsive transit in real-world scenarios with more complex costs, transportation layouts, and passenger behavior, a more sophisticated approach is needed. The proceeding sections will outline a simulation-based approach to comparision and demonstrate that approach using data from the City of Atlanta.

5.2 Comparison of DRT and FRT in Heterogeneous Layouts

This section focuses on extending the fixed-route and demand-responsive transportation comparison techniques to more heterogeneous environments. The methods and examples provided by Quadrifoglio and Li [31] [34] [36], as well as the example in the previous section, deal with homogoneous environments, i.e., the passenger locations and destinations are uniformly spread throught a region, the arrival rates of the passengers are uniform, and the street layouts are grids. In order to better compare demand-responsive and fixed-route transportation in a wider array of settings, the methods must be adapted to handle comparisons from non-uniform environments.

One problem with applying these methods to heterogeneious environments is that closed-form solutions to represent passenger and operator costs are not easily found. Instead, those values must be estimated through simulation. Using the simulator described in Chapter 4, the performance of various transportation systems in heterogeous environments can be accurately measured and compared. In order to simulate in heterogenous environments, four sets of data must be provided.

1. Passenger data: To accurately estimate passenger behavior in a given system, accurage passenger origin and destination data must be available. This data can be estimated through travel demand modeling or collected directly through travel surveys. Much of the passenger data used in this resarch was collected through an Atlanta Regional Commission on-board passenger survey [65].
2. Transit network data: In order to compare systems that use fixed-route transit, accurate schedule data is needed. Schedule data for many agencies is widely available in the General Transit Feed Specification (GTFS) format.
3. Street network data: Street network data can be found through Google Maps, Open Street Maps [38] or a variety of other sources.

4. Demand-responsive transportation layout: For systems that use demand-responsive transportation, the layout and behavior of the demand-responsive transportations system must be defined. For the feeder system studied in this chapter, the demand-responsive coverge area is defined as a radius around each transit station. All passengers with trips beginning or ending within this coverage area will be served by the demand-responsive vehicles in that zone.

Given this set of information, passengers can be entered into the simulator where they will be assigned to the proper fixed-route or demand-responsive network, their complete itineraries will be calculated, and the passenger cost and operator cost of that itinerary will be recorded.

5.2.1 Fixed-Route Transportation Costs

5.2.1.1 Passenger Costs

Passenger costs for fixed-route only transportation are found using the simulation method discussed in Section 4.2.1. For each passenger in the survey data, the passenger's start location, ending location, and request time are passed to the fixed-route transit optimizer. For each passenger the optimizer will determine the fastest route and the passenger's wait, walk, and ride times will be recorded. The total of these times across all the passengers in the simulation provides the total passenger costs for fixed-route transportation.

5.2.1.2 Operator Costs

Operator costs for fixed-route vehicles are found using the general transit feed specification (GTFS) schedules [48]. The GTFS encodes the schedule of all routes operated by a particular agency. These schedules provide precise stop schedules and shape files for each route and trip operated by the agency. From this information, the total number of miles that vehicles travel during a given period and given area can be calculated. The shape files are associasted with both a time and location. Thefore by summing the lengh of each shape file within the geographic study area during the time-period that is being studied, the total distance traveled by the fixed-route fleet can be found. In addition to the total distance traveled by the fleet, the start and end times for each vehicle trip are also provided. This allows for more sophisticated operator cost functions to be considered that not only sum vehicle miles traveled but also operation time. Since a major expense for transit agencies is human resources, knowing how long each driver is required is important for calculating accurate operator costs.

5.2.2 Demand-Responsive Transportation Costs

Passenger costs and operator costs for demand-responsive transportation systems are found through simulation. Using the methods described in Sections 4.2.2 and 4.2.3, the individual cost for each passenger as well as the individual cost for each vehicle is found. As each new passenger enters the system, the passenger is assigned to the optimal vehicle to service that passenger's trip and the route of that vehicle is optimized to minimize the combined cost for all passengers assigned to that vehicle as well as the vehicle cost itself. The precise location of each vehicle is known at each second during the simulation. As new passengers enter the simulation, vehicle costs and passenger costs on continually update to reflect any routing changes made. At the conclusion of the simulation, the total walk, wait and transit times for each passenger are summed as well as the total distance traveled by each vehicle and total operating time for each vehicle. This data provides sufficient information to calculate accurate passenger and operator costs for demand-responsive transportation.

5.3 Comparing DRT and FRT in Atlanta

The simulation-based approach to comparing demand-responsive and fixed-route transit costs is applied to a proposed demand-responsive feeder system in the city of Atlanta, GA. For this exercise, the city of Atlanta is divided into 38 feeder areas or subnets. Each subnet is circular, one mile in radius, and is centered around one of Atlanta's 38 rail stations operated by the Metropolitan Atlanta Rapid Transit Authority (MARTA). These circular feeder areas are reminiscent of the ring-radial structures discussed by Diana in [31], except that the street layout within these feeder areas follow no strict pattern. A map illustrating the locations of these subnets is shown in Figure 5-4.

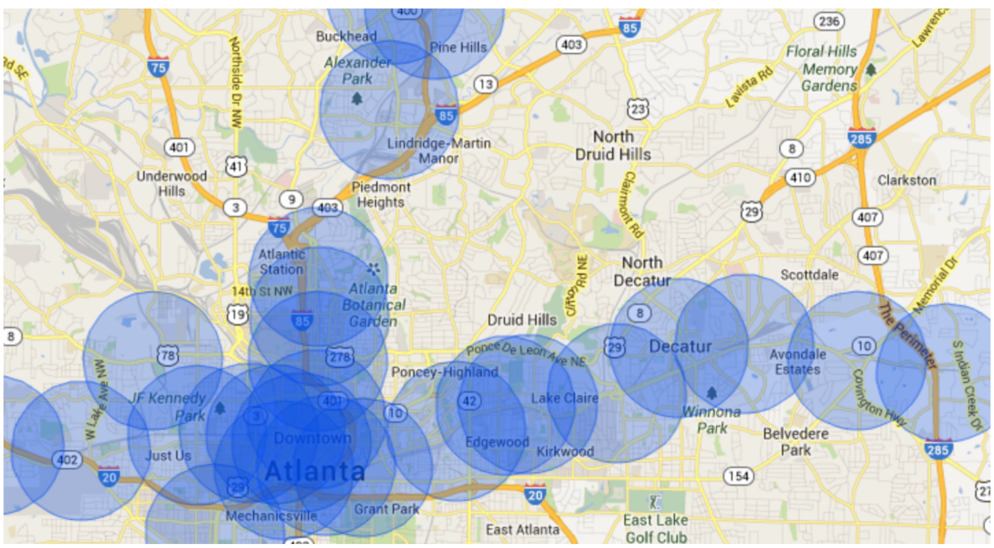

Figure 5-4: Circular subnets centered on each MARTA station.

If a passenger begins his or her trip within one of the subnets, the first leg of his or her trip will be handled by a demand-responsive vehicle that will transport the passenger between the origin location and the nearest rail station. Similarly, if a passenger's trip ends within one of these coverage areas, the final leg of the passenger's trip will be handled by a demand-responsive vehicle. It is important to note that this particular setup is not optimized. It is merely a proposed demand-responsive solution seeking to improve performance for passengers whose trips begin or end near a MARTA rail station. The purpose of this comparison is to determine if implementing this particular demand-responsive framework within each of these zones will improve passenger satisfaction or reduce cost to the transit operator. Choosing an optimal framework is the topic of Chapter 6.

In addition to comparing the performance of this system to the current fixed-route system, the comparison will also determine the level of passenger demand at which the current fixed-route transit system becomes more efficient than the proposed demand-responsive feeder system.

The passenger data used in this study was collected through passenger surveys conducted by the Atlanta Regional Commission [65]. This survey data provides the exact origin and destination of MARTA travelers as well as the time of day they used the system. During the morning peak hours, between 6:00 AM and 9:00 AM, approximately 4,000 riders responded whose destinations and origins lay within one mile of a MARTA station. The goal of this simulation is to determine, for this set of passengers, whether a fixed-route or dynamic route system will best meet the passengers' needs. The dynamic system performance will be analyzed on a city-wide basis as well as on a subnet-by-subnet basis.

The 38 subnets studied cover a wide array of street layouts. Subnets in the city center are often located in gridded street layouts, such as the gridded system near midtown shown in Figure 5-5. Subnets that are farther from the city center become increasingly suburban in design, such a Chamblee station shown in Figure 5-6. This wide array of subnet layout allows for suburban and city layouts to be individually studied and compared.

Figure 5-5: Detail of the Midtown MARTA station with overlapping areas removed.

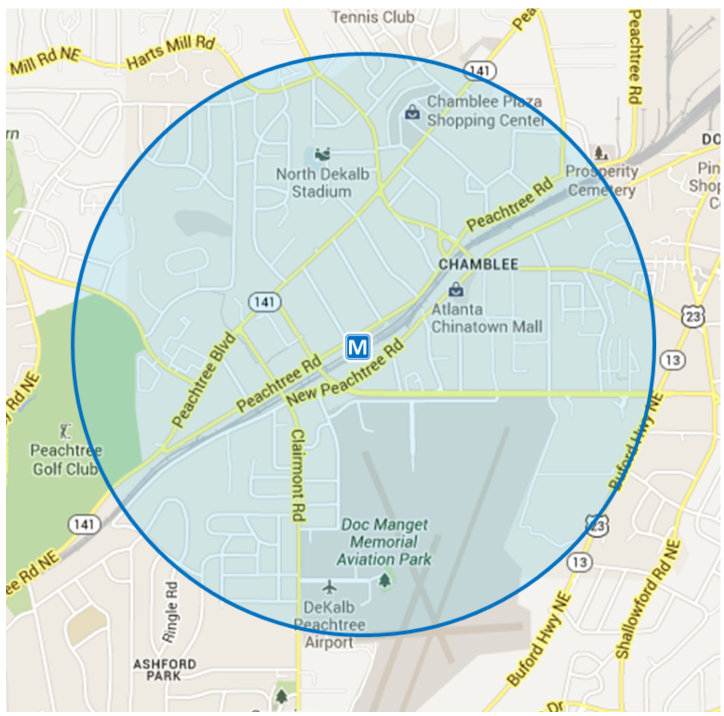

Figure 5-6: Detail of the Chamblee MARTA station with a circular zone.

5.3.1 Creating a Set of Passengers for Study

. The passenger data collected during the survey represents approximately 10% of the MARTA riders. In order to have a set that represents 100% of the riders, the data set must be upsampled. The method of upsampling is as follows.

1. For each passenger origin location, define a circle that is ¼ mile in radius around that location.
2. Within this circle, uniformly create N new origin locations representing N new passengers, where N represents the upsample rate.
3. For each of these new passengers, assign a random departure time weighted to the match the survey departure time statistics. (e.g., if 10% of all departures from the survey occur between 7:30 AM and 8:00 AM, then 10% of all upsampled departures should also occur between these times.)
4. Repeat steps 1 and 2 for each passenger's destination location. After this step a set of origin locations and associated departure times will have been created as well as separate list of destination locations.
5. Randomly assign each location from the origin list to a location from the destination list until all locations have been assigned.

5.3.2 Fixed-Route Transportation Costs in Atlanta

Once the set of passenger origins and destinations is created, calculating fixed-route passenger and operator costs can begin. An OpenTripPlanner instance was created for the city of Atlanta where the street layout was provided by Open Street Map and the GTFS schedule was provided by MARTA. The optimal itineraries for all passengers within one mile of a rail station were calculated by passing each passenger's trip request, one-by-one into the OpenTripPlanner optimizer using the method outlined in Section 4.2.1. For these itineraries, optimality refers to minimum door-to-door passenger travel time. For this set of passengers, the average time to complete a trip was 54 minutes. This 54 minute time is broken down into 11 minutes walking, 16 minutes waiting for transit, and 27 minutes spent riding transit.

The total VMT for the MARTA bus fleet operating within the 1 mile service area around each rail station during this time period is 7,153 miles (11,516 km). The total VMT for the fixed-route service within a given geograhical area and time is derived from the GTFS data, and was found using the method described in Section 5.2. These passenger travel times and vehicle miles traveled will act as the benchmark against which demand-responsive performance will be compared.

5.3.3 Demand-Responsive Transportation Costs in Atlanta

To determine the performance of a demand-responsive feeder system, a simulation was run drawing passengers from the same survey set. However, instead of using a constant passenger arrival rate, the passenger arrival rate was varied from 8 passengers per minute to 200 passengers per minute in order to simulator passenger demand at different times of the day. The various levels of passenger demand were derived by altering the upsample rate described in the previous section For six different passenger demand levels, simulations were run to determine DRT performance. The performance was calculated by simulating optimal vehicle and passenger routes using the methods described in Section 4.2.

The results of these simulations are shown in Figures 5-7 through 5-11. Figure 5-7 shows the vehicle miles traveled of the fixed-route and demand-responsive feeder systems for the six levels of passenger demand. The chart indicates that the DRT feeder system suggested for Atlanta was able to meet the customer demand with fewer vehicle miles traveled for all demand levels below 67 passengers per minute. Above 67 passegners per minute, FRT becomes more efficient in terms of vehicle miles traveled by the fleet.

Figure 5-8 shows the average door-to-door trip times for FRT and the DRT feeder system. The DRT feeder system, on average, was only able to shorten the passengers' trip times for very small demand levels across the city. This chart indicates that this particular DRT setup in Atlanta may have some cost savings for the transit operator, but those savings come at the cost of increased travel times for the passengers. Considering that VMT cost savings only occur during a small portion of the service day and that the total door-to-door passenger travel time is signficantly increased, this particular demand-responsive setup should not be applied to the city as a whole. However, further analysis indicates that portions of the city may be better suited to demand-responsive transportation.

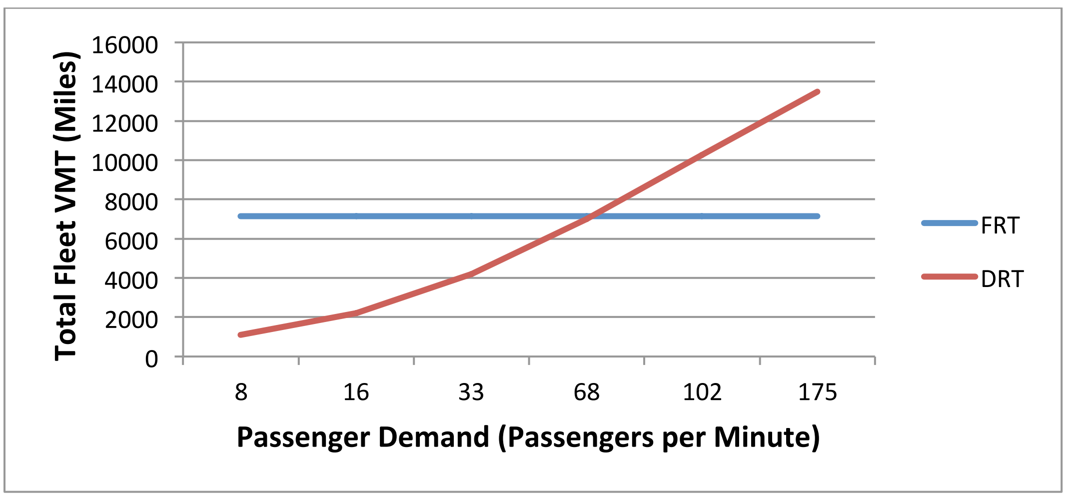

Figure 5-7: Vehicles miles traveled by the fleet in a city-wide simulation.

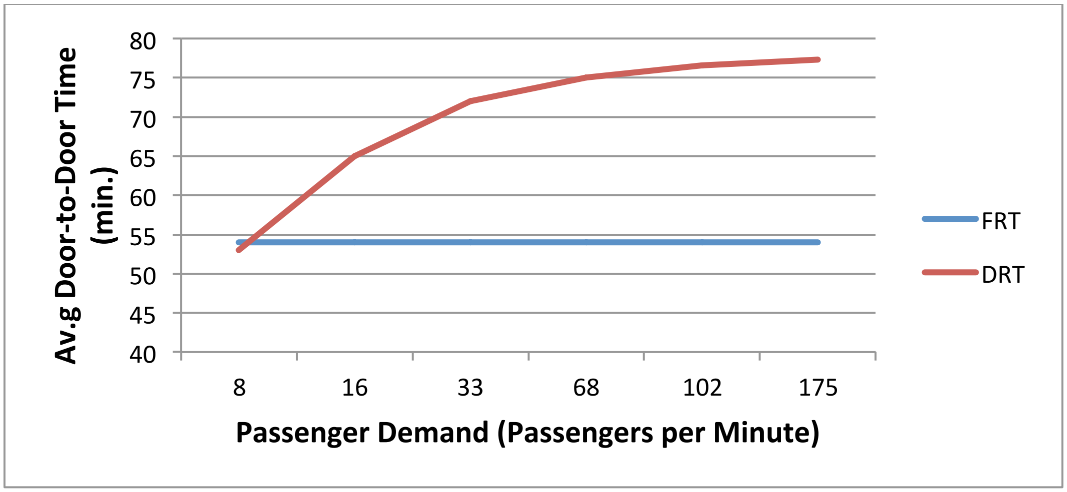

Figure 5-8: Average passenger travel time in a city-wide simulation.

One area where the DRT feeder system compares better with the FRT system is the subnet centered on the Chamblee MARTA station. The Chamblee Station subnet is far from the city center and has a suburban street layout. Figures 5-9 through 5-11 show the DRT and FRT comparisons for vehicles operating in this subnet and for passenger travel times whose trips either ended or began within this subnet. Figure 5-9 shows that the DRT setup in this region has signficant VMT savings for all passenger demand levels tested. Unlike the city-wide travel times shown in Figure 5-8, the DRT system decreased travel times for passengers for all demand levels below 44 passengers per minute. This is shown in Figure 5-10. After normalizing the VMT and passenger travel time costs, shown in Figure 5-11, the point at which FRT begins outperforming DRT is shown to be 165 passengers per minute. This is signficantly higher than the FRT break even point for the city-wide scenario.

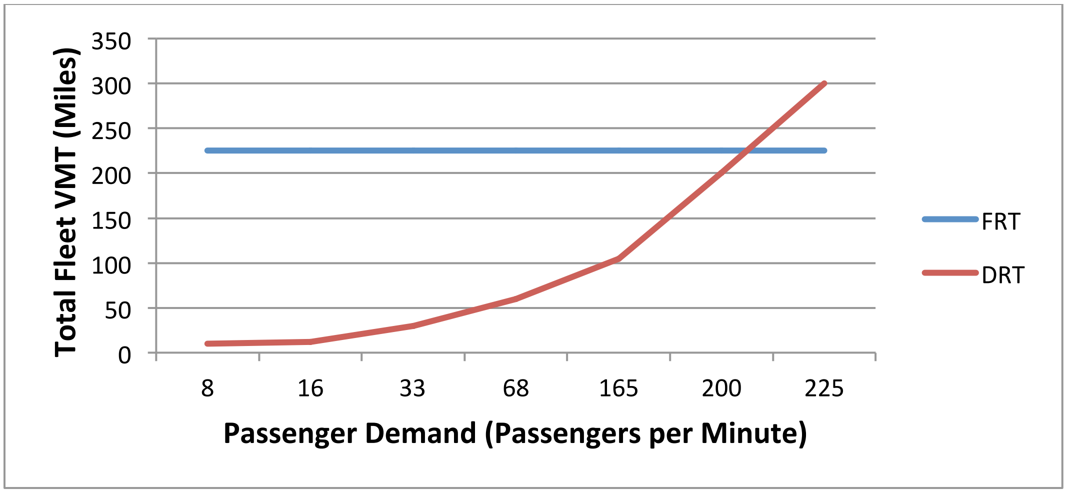

Figure 5-9: Vehicles miles traveled in the Chamblee Station Subnet.

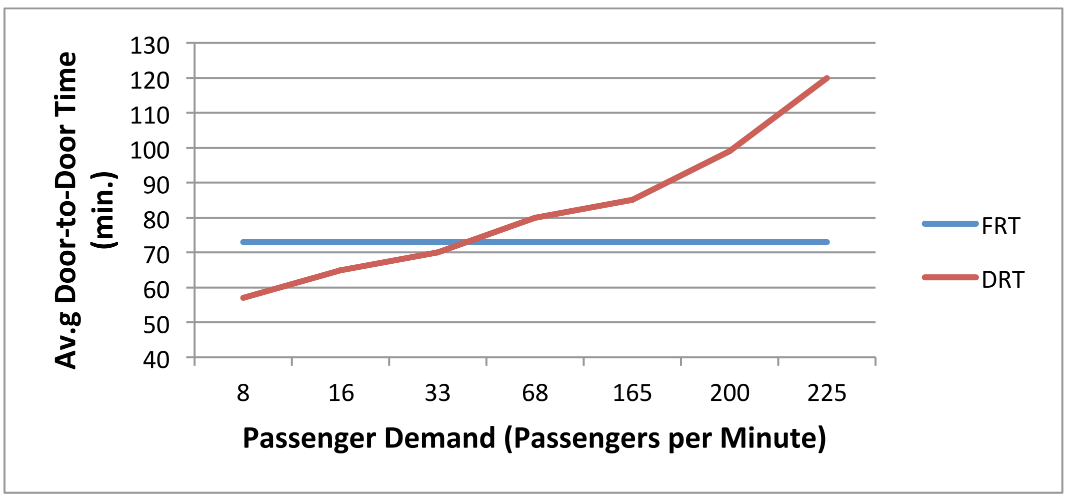

Figure 5-10: Average passenger travel time in the Chamblee Station Subnet.

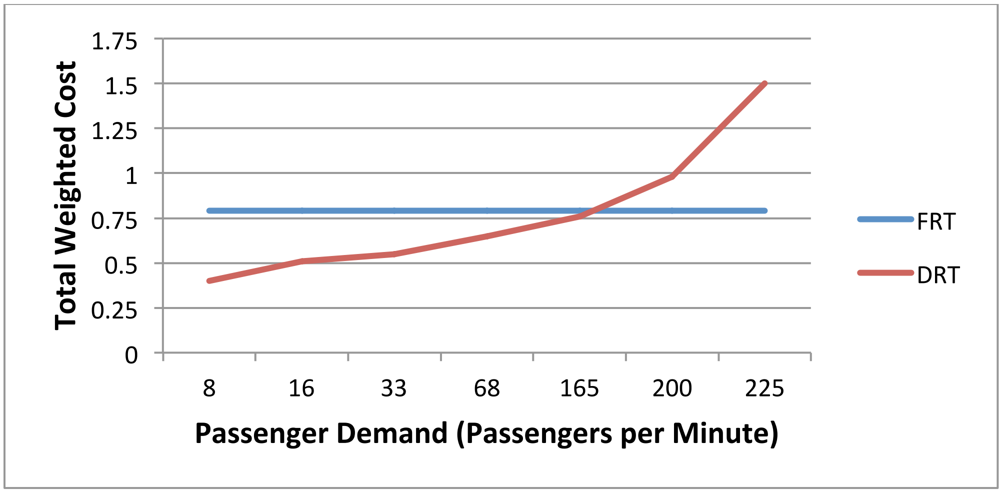

Figure 5-11: Combined operator and passenger costs in the Chamblee Area Subnet.

These results show that DRT could provide a less expensive alternative for handling trip requests for stations with relatively low demand at off-peak hours. Other stations such as North Springs Station, Doraville Station, and Brookhaven Station also showed similar improvements in costs by using DRT. All of these stations have suburban street layouts. What is interesting to note here is that demand-responsive transit in these suburban areas was less efficient, in terms of operation cost and travel times, than demand-responsive transit in the more urban areas. However, fixed-route transit is also less efficient in these suburban areas. For instance, in the Chamblee subnet, the average fixed-route door-to-door time was 72 minutes, the system average was 54 minutes. The fact that DRT was more efficient than FRT in these areas does not necessarily mean that the DRT feeder system is best-suited for suburban areas. Instead it suggests that DRT is more adept at handling the difficulties of servicing a low-density, suburban area than a FRT system.

This method of comparison also brings to light areas that need improvement in the DRT algorithms. For instance in this simulation, the demand-responsive algorithm consistently reduced vehicle miles traveled compared to the fixed-route system.

Unfortunately, this savings to the operator came a large expense to the passengers in terms of convenience. Adjusting the algorithm to give more passenger-friendly route selection could improve the overall performance of the system.

This chapter demonstrated a method of using passenger data, open source mapping and trip optimization tools, as well as the demand-responsive trip simulator to compare fixed-route service and demand-responsive transportation in heterogeneous environments. What is missing is a method of selecting the design of the demand-responsive feeder zones. In this chapter a demand-responsive framework was proposed and tested against the existing fixed-route service. It was determined that for the vast majority of instances, this demand–responsive setup would be inefficient for the City of Atlanta to implement. The next chapter improves upon this proposal by introducing a method of optimizing the demand-responsive framework that builds upon the comparison techniques used in this chapter.

CHAPTER 6

OPTIMAL DESIGN OF DEMAND-RESPONSIVE FEEDER ZONES

6.1 Motivation

As discussed in Section 2.4 of the literature review, there is a need for selecting the optimal size and shape of demand-responsive feeder zones. Chapter 5 introduced a method of comparing fixed-route and demand-responsive transportation performance in heterogeneous environments. The test cases that were presented in Chapter 5 showed mixed results for implementing a demand-responsive transportation system for the city of Atlanta, GA. One problem with the proposed demand-responsive system is that no optimization was performed in order to determine what size and shape each feeder area should be. This chapter will present two improvements to the system proposed in the previous chapter. The first improvement will leverage the comparison technique for heterogeneous environments to build a particle swarm-based optimization algorithm. The feeder zones will be proposed, tested, improved, and tested again until a near-optimal solution is found. The second improvement is to use isochrones to define the demand-responsive coverage areas instead of geometric shapes. In this chapter, the isochrone technique and particle-swarm optimizer are introduced and test cases are presented to demonstrate their efficacy.

Similar to the existing comparison methods, much of the previous research in feeder zone design makes assumptions about street layouts, passenger arrival rates, and passenger locations that do not apply to cities with irregular street layouts and infrequent transit service. Specifically, passenger arrival rates and distribution are often assumed to be uniform and streets are often modeled as gridded or ring-radial systems. Given these assumptions, the practice of designing a feeder zone has often entailed selecting an optimal size of a basic shape of distribution, e.g., square, circle, or hexagon, as well as selecting an optimal number of sub-zones in which to divide the feeder zone. These methods result in concise analytical models for determining optimal demand-responsive zone size and shape for a given passenger demand. The drawback to these methods is that the results do not provide reliable results for suburban areas and do not take into account the nuanced street patterns, transit schedule, or passenger behavior of those areas.

This chapter provides an alternative to these methods that takes advantage of a recent trend of opening transit software and data to researchers. It has become common practice for transit operators to provide schedule data in the form of a general transit feed specification (GTFS) archive. Complementing the open-source release of transit data, is a set of open-source planning tools, such as OpenTripPlanner and the Open Source Routing Machine, that allow researchers to perform thousands of automobile and transit route optimizations very quickly. The demand-responsive zone optimization method proposed and analyzed in this chapter utilizes this data and software tools to create a generic approach that takes into account any street network, transit network, and passenger distribution.

This chapter will introduce and explain a generic demand-responsive zone optimization method. The method will then be used to optimize feeder zone size and shape for locations in the city of Atlanta. The zone size and shape found by this method will be compared to those found with current optimization methods.

The use of basic shapes, e.g., rectangles, circles, or hexagons, are common representations for the coverage areas and walksheds of transit stations and bus stops [66]. Assuming coverage areas with basic shapes allows for analytical solutions to be found that represent the distribution of passengers within the area, arrival rates for passengers within the area, and average walking, waiting, and transit times for those passengers. The drawback of using these shapes is that the actual street network is not accurately taken into account when considering a walking time or driving time. In Figure 6-1, Walker illustrates this shortcoming. In the image on the left, a bus stop with a suburban street layout is shown. In the figure on the right, a stop in a gridded street pattern is shown. These images illustrate that the areas that can be reached by walking from a given point are not accurately represented by a circle. The darkened portion of the streets shows the areas that a passenger can actually reach within a given time. In the suburban layout, the outline of the stop walkshed is of no discernible shape. Coverage areas in suburban areas cannot be easily represented by a simple shape and analytical representations of the driving time or walking times cannot be found. The walkshed of the stop within a gridded street pattern is a much more uniform shape, specifically that of a square rotated 45° with respect to gridded street system.

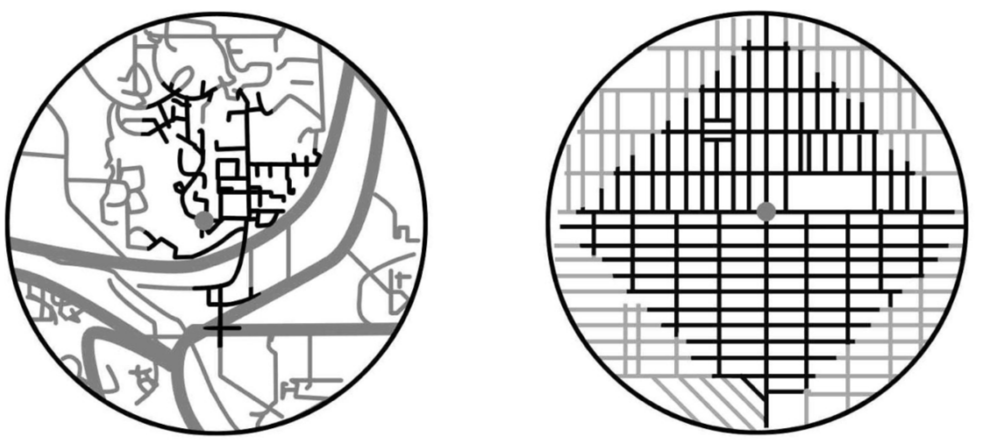

Figure 6-1: Illustration of the area reachable by foot from a transit station during a set period of time [66].

Neighborhoods with a grid layout are more easily represented with analytical formulae than those neighborhoods with suburban layouts. For instance, the optimization example discussed in the literature review for El Cenizo, TX closely matches with a gridded street layout [33], which allows for analytical solutions to be used to define optimal feeder designs. However, as more perturbations are inserted into the grid, such as highways, parks, or irregularly-shaped blocks, the coverage becomes less uniform. In this particular case, Walker is illustrating that assuming basic shapes does not always accurately describe a pedestrian's reachable area, however this same concept can also be applied to motor vehicles. Given a preset starting point, the area reachable by an automobile is unlikely to be a circle or any other basic shape, especially in areas with irregular street patterns.

The inspiration for this research and the network-inspired transportation system is the city of Atlanta, which contains a mix of gridded street layouts as well as irregular street layouts. A detail of a map of Atlanta showing the Chamblee MARTA station and the surrounding area is shown in Figure 6-2. The area around the MARTA station at Chamblee has an irregular, suburban layout. Existing optimization methods for determining the optimal demand-responsive zone layout are unlikely to provide efficient results due the inability of a basic shape to represent this type of street layout.

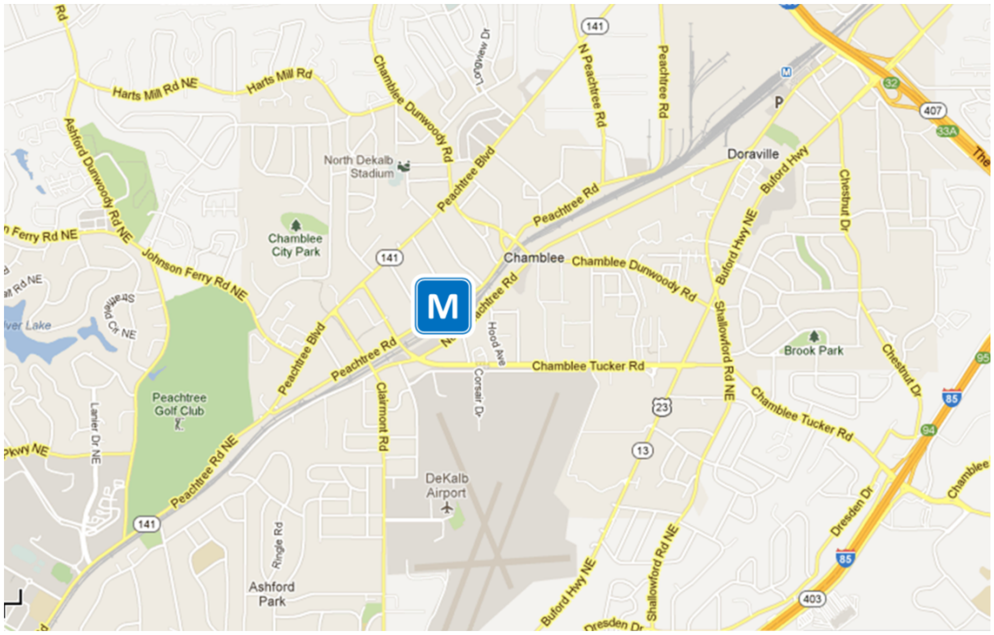

Figure 6-2: Detail of area surrounding Chamblee MARTA Station in Atlanta, GA.

Figure 6-3 shows a rectangular feeder system overlaying the area surrounding the Chamblee MARTA station. Assuming a gridded street system and uniform walking and driving times, implies that the time required to reach Point A from the MARTA station is the same as the time to reach Point B from the MARTA station. This assumption is often utilized to develop analytical solutions for passenger and vehicle travel times as well as to find the optimal size of this coverage area [37]. However, this assumption does not hold up in this particular example. The travel time of the northern-most route is 10 minutes, and the travel time of the southern-most is 8 minutes.

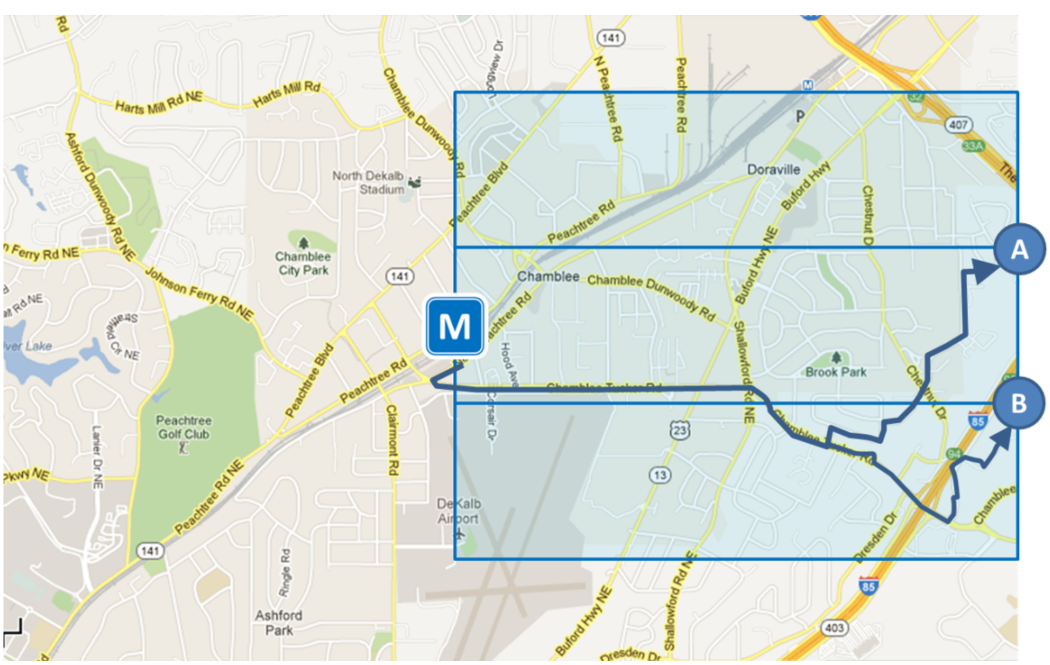

Figure 6-3: Detail of Chamblee MARTA Station with rectangular overlay and two routes shown.

Similarly, assuming a circular radius for walking and driving times does not hold up in an irregular street pattern. In a circular or ring-radial system, all points along the circle are assumed to have nearly identical travel times from the center. Furthermore, travel time between any point within the circle to the center, is assumed to be less than the travel time between a point on the circle to the center. Figure 6-4 shows a circular subnet overlaying the Chamblee MARTA station with two points selected. The travel time between point A on the radius of the circle is 3 minutes, and the travel time from point B, which is within the circle, to the station is 7 minutes. For areas with suburban layouts, a different approach for designing the feeder zones is required.

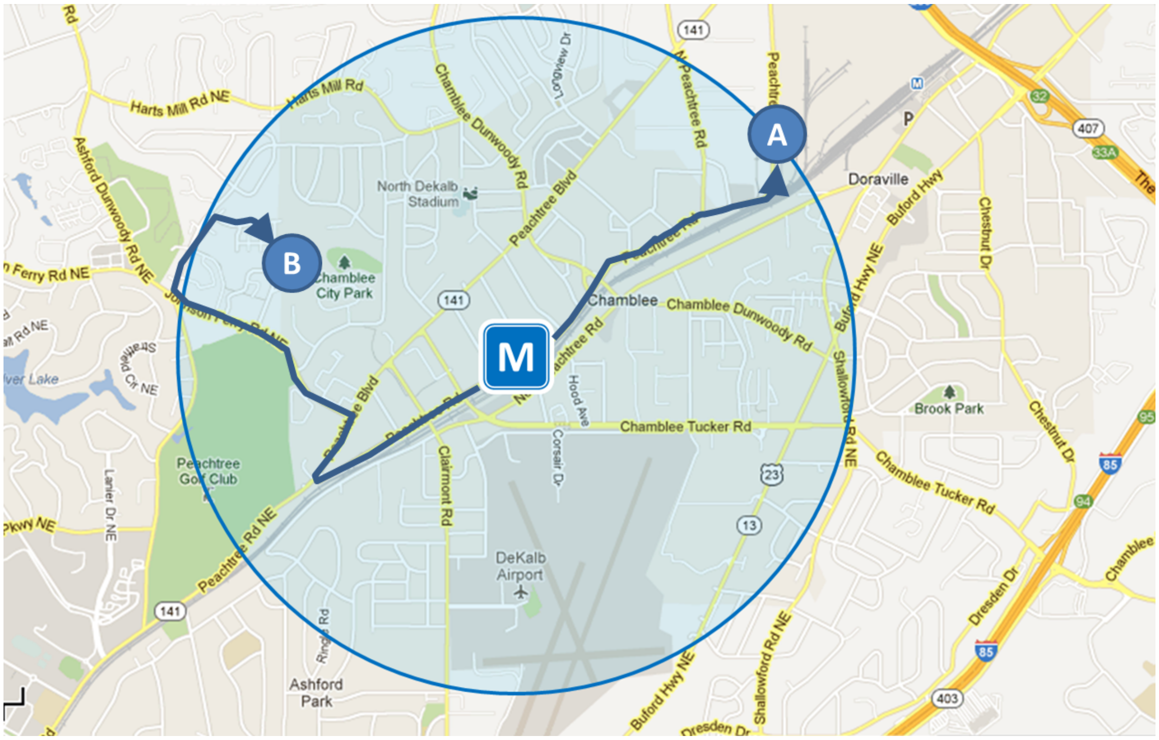

Figure 6-4: Detail of Chamblee MARTA Station with circular overlay and two routes depicted.

In order to better represent irregular street networks, this research proposes using isochrones as the foundation for deciding the size and shape of demand-responsive service areas. An isochrone is a contour line where the travel times between any two points along the line and a reference point are equal. For instance, Figure 6-5 shows an isochrone where the Chamblee MARTA station acts as the reference point. Driving between the MARTA station and any point along that line will take exactly five minutes.

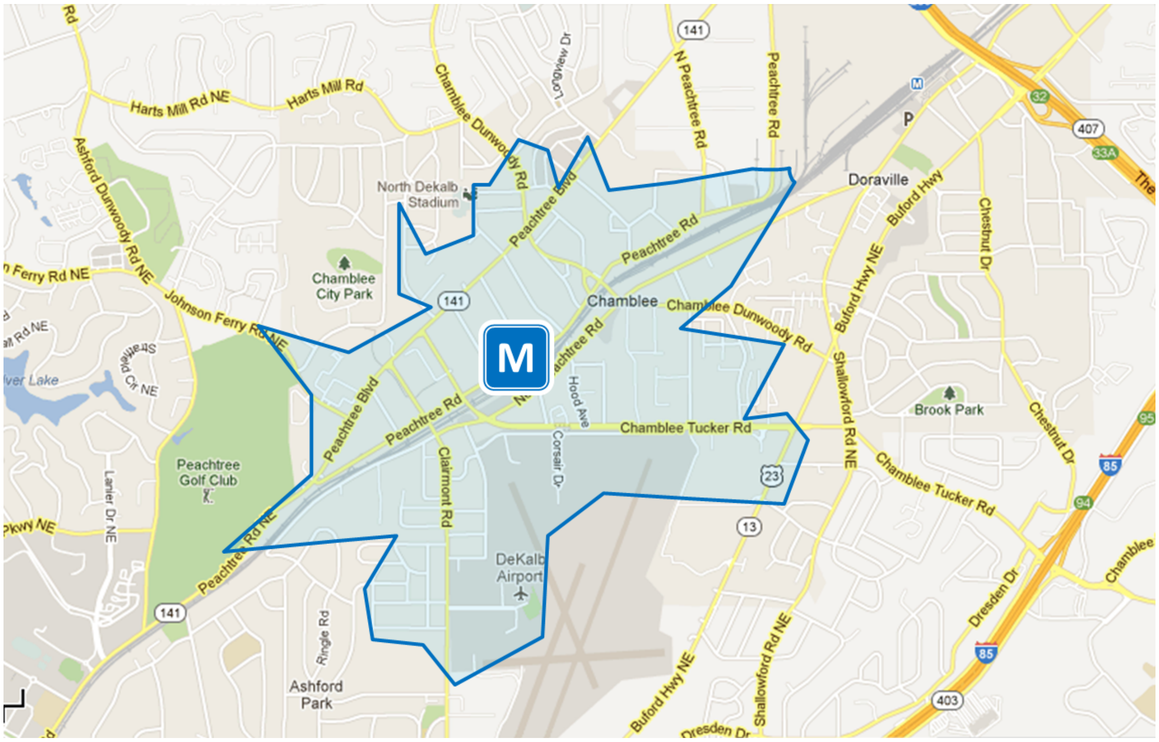

Figure 6-5: Detail of Chamblee MARTA Station with a driving isochrone depicted.

Letting an isochrone decide the shape of a feeder zone allows for the nuances of an irregular street network to be considered. Driving time isochrones can consider not only straight-line distance but also driving-distance, driving speed, intersections, and even traffic conditions. Since the street network dictates the shape of the feeder zone, the only decision left to the transportation engineer is to decide the optimal size of the feeder zone. Determining the size of the feeder zone is a main goal of this research, and the process is described in detail in the following section.

In addition to implanting a boundary based on driving time, a second boundary based on walking time can also be implemented to improve performance. The second boundary will define an inner boundary for the feeder zone. While the outer boundary represents the maximum distance that a vehicle will travel to pick up a passenger, the inner boundary will define the maximum walking distance that a passenger is expected to walk without requiring a demand-responsive vehicle. Figure 6-6 shows an example of both inner and outer isochrones for Chamblee Station. Only passengers between the two isochrones will be eligible for a demand-responsive trip. Passengers outside of the larger isochrone are not within the service area, and passengers inside the smaller isochrone will be required to walk. The theory behind this is that a large number of trips originate or end within a short distance of the rail station and requiring demand-responsive vehicles to service these short trips will require more time than if the passengers walked those distances. This theory will be tested and the optimal walking distance for a given station is found in Section 6.4.

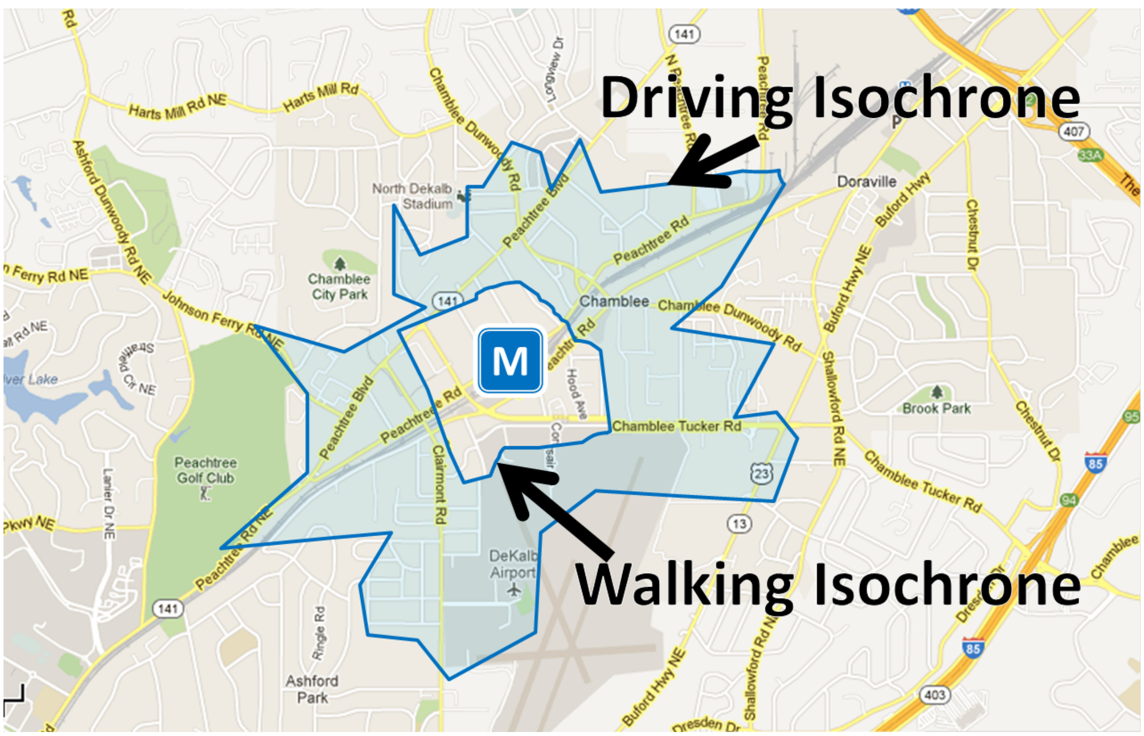

Figure 6-6: Detail of Chamblee MARTA Station with walking and driving isochrones depicted.

An additional benefit of utilizing isochrones in feeder area design is identifying boundaries between feeder zones. In order to provide efficient service, it is desirable to cover as much area as possible with service without duplicating services with overlapping feeder areas [66]**.** An image from chapter 5 is reproduced below in Figure 6-7. In this image, each rail station in the MARTA system is shown with a radius of one mile.

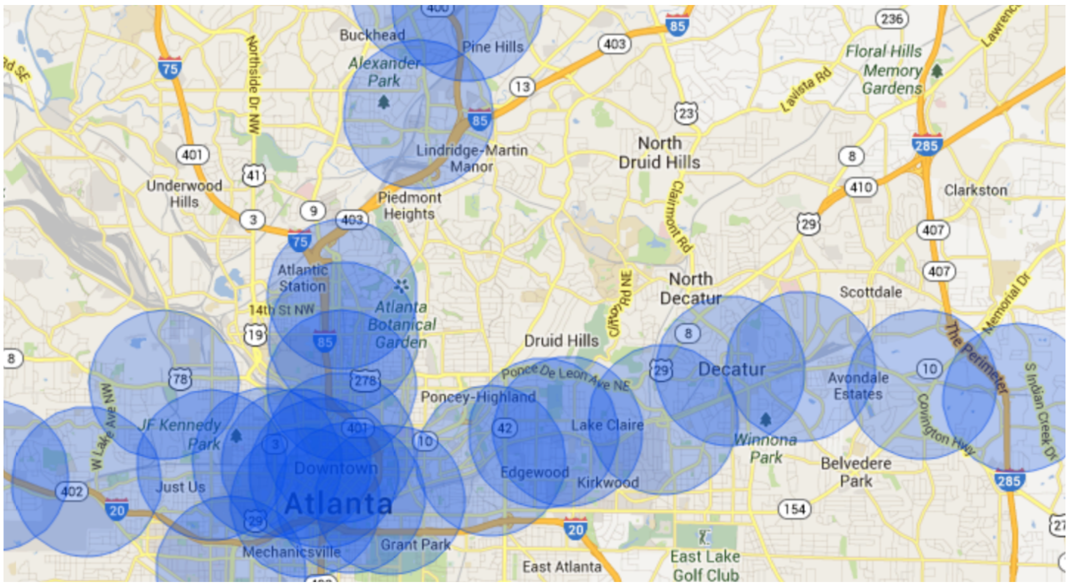

Figure 6-7: MARTA stations shown with a one-mile radius around each one.

Simulations in Chapter 5, assumed this layout as the basis for comparing demand-responsive transportation with fixed-route transportation. In this image, large overlaps between feeder zones are present as well as large swaths of land that are not covered with demand-responsive transportation. Using isochrones to determine the coverage area for each MARTA station will identify where the boundary between each station should begin and end, in terms of driving time, as well as identify more accurate boundaries for how far vehicles can efficiently travel away from a gateway. Figure 6-8 shows the same

MARTA rail stations, except this time instead of an arbitrary one-mile radius around each station, the service areas are defined by driving times.

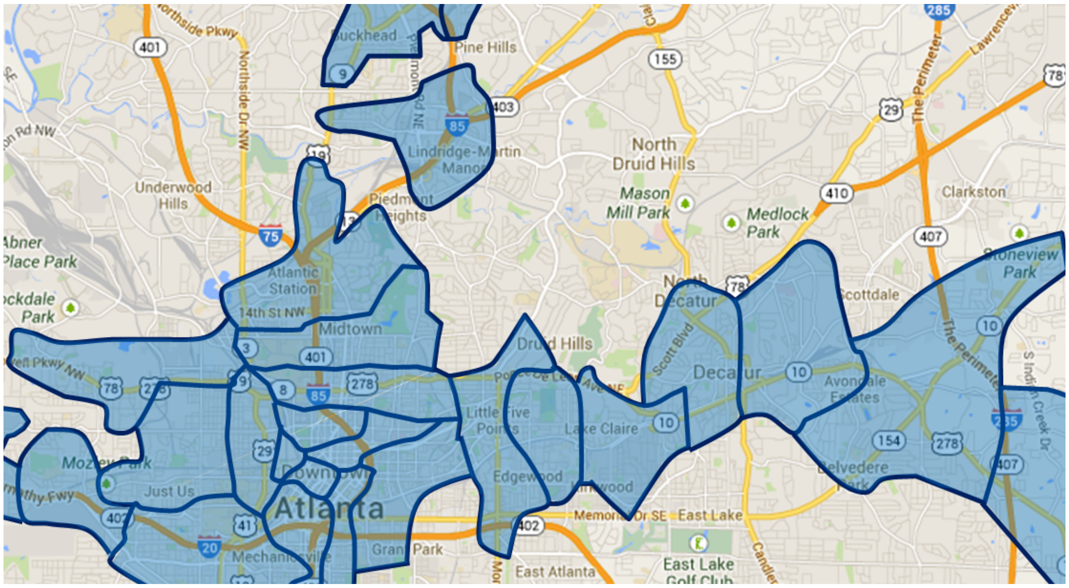

Figure 6-8: MARTA stations shown with time-based coverage-area boundaries.

Each boundary represents either a boundary between the service area of two stations or the point at which a vehicle is more than five minutes away from the nearest rail station. The use of isochrones and driving time averages results in a map with a realistic representation of how far vehicles can travel in a given time. The exact size of these coverage areas can be further refined using the particle swarm optimization technique described in section 6.3.

6.2 Defining the Optimal Zone Size

Allowing the driving and walking isochrones to dictate the shape of the feeder zone leaves only two variables left to optimize, the size of the driving isochrone, and the size of the walking isochrone. Just as Chapter 5 used a simulation-based approach to compare demand-responsive and fixed-route transportation, a simulation approach will also be used to determine the optimal sizes of the isochrones.

Before the optimization process can begin, the optimization fitness function, or cost function, must be defined. Equations 6.1-6.13 will step through the process of determining the net cost of implementing a demand-responsive feeder system to replace a fixed-route system. The costs represented in these equations are not meant to incorporate every feasible cost associated with operating a transit agency. Instead, they are meant to represent the variable costs that can be controlled in real time by optimizing vehicle routes and passenger itineraries.

The total fitness or cost (J_T) of operating a transportation system is a weighted combination of operator costs (J_O) and passenger costs (J_P), show in Equation 6.1.

$$J_T = \alpha J_O + \beta J_P \tag{6.1}$$

The operator costs can be further reduced to the combination of the costs for operating demand-responsive service (J_D) and the cost for operating the fixed-route portion of the service (J_f).

$$J_O = \alpha_D J_D + \alpha_f J_f \tag{6.2}$$

The cost of operating demand-responsive transportation is the sum of the cost of operating each demand-responsive vehicle. Similarly, the total cost of operating fixed-route transit is the sum of the cost of operating each fixed-route vehicle. These costs are shown in Equations 6.3 and 6.4 where $p_{d,i}$ is the total cost of operating demand-responsive vehicle i, and $p_{f,i}$ is the total cost of operating fixed-route vehicle i. M_d is the total number of demand-responsive vehicles in the system, and M_f is the total number of fixed-route vehicles in the system.

$$J_D = \sum_{i=1}^{M_d} v_{d,i} \tag{6.3}$$

$$J_f = \sum_{i=1}^{M_f} v_{f,i} \tag{6.4}$$

The total passenger cost is the sum of each individual passenger's cost, which can be broken down into wait, walk, and transit times.

$$J_P = \sum_{i=1}^{N} p_i = \sum_{i=1}^{N} (w_{wk}\tau_{wk} + w_{wt}\tau_{wt} + w_{tr}\tau_{tr}) \tag{6.5}$$

The total cost of operating an integrated demand-responsive system includes the total operational and passenger costs within the fixed-route portion of the system and the demand-responsive portion of the system. The total cost of a fixed-route system includes no demand-responsive vehicle costs. However, in order to compare a fixed-route system with a demand-responsive system it helps to divide the fixed-route system into two portions: the portion that will be replaced by demand-responsive vehicles and the protion that will not be replaced by demand-responsive vehicles.

For example, in a fixed-route transportation system, the set of fixed-route vehicle costs can be represented as

$$V_f = [v_{f,1}, v_{f,2}, \dots, v_{f,M_f}],$$

and the set of passenger costs can be represented as

$$P = [p_{f,1}, p_{f,2}, \dots, p_{f,N}].$$

To simplify the problem of comparing costs between fixed-route and demand-responsive systems, the vehicles costs can be rewritten as:

$$V_f = [v_{f,1}, v_{f,2}, \dots, v_{f,M_I,} v_{f,M_I+1}, \dots, v_{f,M_I+M_O}],$$

and the passenger costs can be rewritten as :

$$P = [p_{f,1}, p_{f,2}, \dots, p_{f,N_I}, p_{f,N_I+1}, \dots, p_{f,N_I+N_O}].$$

Where M_I is the number of fixed-route vehicles operating inside the demand-responsive coverage that will be replaced by demand-responsive vehicles, and M_O is the number of fixed-route vehicles operating outside the demand-responsive coverage area. These routes will not be replaced. Similarly N_I is the number of passengers with trips within the coverage area, and N_O is the number of passengers that do not fall within the coverage area.

Using this nomenclature, the total cost of operating an integrated demand-responsive transportation system can be represented by Equation 6.6.

$$J_T = \alpha\left(\alpha_D \sum_{i=1}^{M_d} v_{d,i} + \alpha_f \sum_{i=M_I+1}^{M_I+M_O} v_{f,i}\right) + \beta\left(\sum_{i=1}^{N_I} p_{d,i} + \sum_{i=N_I+1}^{N_I+N_O} p_{f,i}\right) \quad (6.6)$$

Table 6-1: Variable explanations for passenger and operator net cost functions.

Variable	Definition
$\alpha_D \sum_{i=1}^{M_d} v_{d,i}$	The total cost of all demand responsive vehicles used in the system multiplied by a weighting factor α_D.
$\alpha_f \sum_{i=M_I+1}^{M_I+M_O} v_{f,i}$	The total cost of all fixed-route vehicles that were not replaced with demand-responsive vehicles multiplied by a weighting factor α_f.
$\sum_{i=1}^{N_I} p_{d,i}$	The total passenger costs for all passengers with at least one demand-responsive leg.
$\sum_{i=N_I+1}^{N_I+N_O} p_{f,i}$	The total passenger costs for all passengers using fixed-route only.

The cost of operating a purely fixed-route system is a combination of operating all the fixed-route vehicles as well as the cost incurred by every passenger on those vehicles.

$$J_{T,frt} = \alpha\left(\alpha_f \sum_{i=1}^{M_I+M_O} v_{f,i}\right) + \beta \sum_{i=1}^{N_I+N_O} p_{f,i} \tag{6.7}$$

The demand-responsive optimizer seeks to find the minimum net cost of operating demand-responsive transportation. Shown in 6.8, the net total cost of switching from a fully fixed-route system to in integrated demand-responsive system is the total cost of operating an integrated demand-responsive system minus the cost of operating a fixed-route only system.

$$J_{T,net} = J_T - J_{T,frt} \tag{6.8}$$

$$J_{T,net} = \alpha\left(\alpha_D \sum_{i=1}^{M_d} v_{d,i} + \alpha_f \sum_{i=M_I+1}^{M_I+M_O} v_{f,i}\right) + \beta\left(\sum_{i=1}^{N_I} p_{d,i} + \sum_{i=N_I+1}^{N_I+N_O} p_{f,i}\right) - \left[\alpha\left(\alpha_f \sum_{i=1}^{M_I+M_O} v_{f,i}\right) + \beta \sum_{i=1}^{N_I+N_O} p_{f,i}\right] \tag{6.9}$$

Equation 6.9 can be rearranged as equation 6.10.

$$J_{T,net} = \alpha\left(\alpha_D \sum_{i=1}^{M_d} v_{d,i} + \alpha_f \left(\sum_{i=M_I+1}^{M_I+M_O} v_{f,i} - \sum_{i=1}^{M_I+M_O} v_{f,i}\right)\right) + \beta\left(\sum_{i=1}^{N_I} p_{d,i} + \sum_{i=N_I+1}^{N_I+N_O} p_{f,i} - \sum_{i=1}^{N_I+N_O} p_{f,i}\right) \tag{6.10}$$

The cost incurred by fixed-route only passengers traveling within the demand-responsive zone is shown in 6.11.

$$\sum_{i=1}^{N_I} p_{f,i} = \sum_{i=1}^{N_I+N_O} p_{f,i} - \sum_{i=N_I+1}^{N_I+N_O} p_{f,i} \tag{6.11}$$

The cost incurred by fixed-route vehicles operating within the demand-responsive zone is shown in 6.12.

$$\sum_{i=1}^{M_I} v_{f,i} = \sum_{i=1}^{M_I+M_O} v_{f,i} - \sum_{i=M_I+1}^{M_I+M_O} v_{f,i} \tag{6.12}$$

Inserting 6.11 and 6.12 into 6.10 yields 6.13. This is the total net cost of switching from a fixed-route only transit system to an integrated demand-responsive system. This equation is what the particle swarm optimizer attempts to minimize.

$$J_{T,net} = \alpha\left(\alpha_D \sum_{i=1}^{M_d} v_{d,i} - \alpha_f \sum_{i=1}^{M_I} v_{f,i}\right) + \beta\left(\sum_{i=1}^{N_I} p_{d,i} - \sum_{i=1}^{N_I} p_{f,i}\right) \tag{6.13}$$

6.3 Particle Swarm Optimization

A drawback of taking a simulation-based approach to finding optimal zone size is that optimizing vehicle routes for dozens of vehicles and hundreds or even thousands of passengers is a time-consuming process. Figure 6-9 shows processing times for 20 trial simulations of operating demand-responsive transportation at Midtown Station in Atlanta, GA on a Sunday afternoon. Each trial simulated a three-hour period from 11AM until 2PM for a variety of feeder zone boundaries. For large boundaries, more trips were simulated and optimized, leading to higher computation times. There is a clear correlation between the number of passengers requesting trips and the processing time required to simulate these trips. The average time required to simulate each of these three-hour trials was 151 minutes.

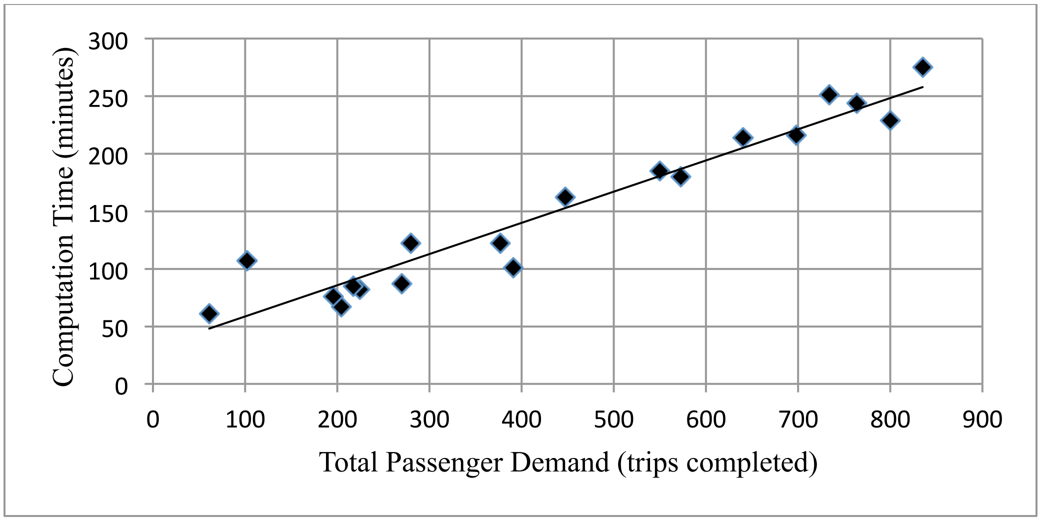

Figure 6-9: Simulation processing times on an Intel Core i7-740QM with 8,193 MB of RAM versus passenger trips completed during the simulation.

Each time a new isochrone is tested, a new simulation must be run. When attempting to find the optimal isochrone sizes, many such simulations must be run, which can lead to hours or even days of processing time to locate the optimal isochrone. For the twenty simulations represented in the chart above, the total simulation time was 50 hours and 20 minutes. In order to mitigate these long simulation times, a smart optimization routine is needed to minimize the number of simulations that must be run in order to find the optimal isochrone sizes.

This research uses a particle swarm optimization routine to minimize the number of samples in order to find a near-optimal solution. A particle swarm optimizer was chosen because it can efficiently search a 2-D search space, and quickly converge on a near-optimal solution without becoming trapped in local optima [67]. The 2-D search space in this application consists of finding the optimal walking isochrone size and optimal driving isochrone size. Each of these is measured in seconds. Alternatively, if a basic shape, such as a circle, is used to define the coverage area the 2-D search space can represent the radii of the inner and outer circles. A pseudo-code representation of the particle swarm algorithm used in this work is shown in Figure 10. Table 6-2 provides a definition of the variables used. The algorithm initializes a small set of particles. For the application presented here, 6-10 particles are used, and are initialized uniformly across the search space. The range of the search space varies with the area under study. While the minimum boundaries for walking time or driving time can never be less than zero seconds, the maximum size has no hard boundary. In the simulation examples below, an explanation for how the maximum boundary sizes were chosen for each example will be explained. After initialization, the net passenger and operator cost for each particle is found and the values and velocities of each particle are updated as described in Figure 6-10.

.

$J_{g,min} = inf$

for m=0…M: //Initialize all particles uniformly across the search space

  $\begin{bmatrix} x_{1,0} \\ x_{2,0} \end{bmatrix}_m \sim \mathrm{U} \begin{bmatrix} x_{1,min}, x_{1,max} \\ x_{2,min}, x_{2,max} \end{bmatrix}$

  $J_{m,min} = inf$

  $\begin{bmatrix} v_{1,0} \\ v_{2,0} \end{bmatrix}_m = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

for n=0…N: //For each particle at each step, find the cost and update the particle

  for m=0…M:

    $J_{m,n} = f\left(\begin{bmatrix} x_{1,n} \\ x_{2,n} \end{bmatrix}_m \right)$ //Find the cost for this a particle

    if $J_{m,n} < J_{m,min}$: //Check to see if a minimum for this particle is found

      $J_{m,min} = J_{m,n}$

      $\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix}_m = \begin{bmatrix} x_{1,n} \\ x_{2,n} \end{bmatrix}_m$

    if $J_{m,n} < J_{g,min}$: //Check to see if a global minimum is found

      $J_{g,min} = J_{m,n}$

      $\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix}_g = \begin{bmatrix} x_{1,n} \\ x_{2,n} \end{bmatrix}_m$

    $r_{1,n}, r_{2,n} \sim \mathrm{U}[0,1]$

    $\begin{bmatrix} v_{1,n+1} \\ v_{2,n+1} \end{bmatrix}_m = \begin{bmatrix} v_{1,n} \\ v_{2,n} \end{bmatrix}_m + c_1 r_{1,n} \left[\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix}_m - \begin{bmatrix} x_{1,n} \\ x_{2,n} \end{bmatrix}_m \right] + c_2 r_{2,n} \left[\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix}_g - \begin{bmatrix} x_{1,n} \\ x_{2,n} \end{bmatrix}_m \right]$

    $\begin{bmatrix} x_{1,n+1} \\ x_{2,n+1} \end{bmatrix}_m = \begin{bmatrix} x_{1,n} \\ x_{2,n} \end{bmatrix}_m + \begin{bmatrix} v_{1,n+1} \\ v_{2,n+1} \end{bmatrix}_m$ //Update the particle for the next step

Figure 6-10: Pseudo-code representation of the particle swarm optimization algorithm used to find optimal feeder zone boundaries.

Table 6-2: Variable definitions for the Particle Swarm Optimizer pseudo-code.

Variable	Definition
$x_{1,n}$	Outer isochrone size at step *n*.
$x_{2,n}$	Inner isochrone size at step *n*.
$\begin{bmatrix} x_{1,n} \\ x_{2,n} \end{bmatrix}_m$	Values of particle *m* at step *n*.
$\begin{bmatrix} v_{1,0} \\ v_{2,0} \end{bmatrix}_m$	Velocities of particle *m* at step *n*.
M	Total number of particles.
N	Total number of steps.
$\begin{bmatrix} x_{1,min}, x_{1,max} \\ x_{2,min}, x_{2,max} \end{bmatrix}$	Maximum and minimum values for inner and outer isochrones.
$J_{m,n}$	Cost of particle *m* at time *n*.
$J_{m,min}$	Minimum cost of particle *m*.
$\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix}_m$	Value of particle *m* with lowest cost.
$J_{g,min}$	Minimum cost of all particles.
$\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix}_g$	Value of particle with lowest cost across all particles.
c_1, c_2	Operator-chosen acceleration variables.
$r_{1,n}, r_{2,n}$	Random acceleration variables.

6.4 Feeder Zone Test Cases in Atlanta

6.4.1 Description of Feeder Zone Operation and Costs

To assess the effectiveness utilizing particle swarm optimization to find optimal isochrone sizes for demand-responsive feeder zones, a series of simulations were run. These simulations attempt to illustrate the effectiveness of utilizing isochrones as boundaries for feeder zones by comparing feeder zones defined by isochrones to feeder zones defined by straight line distances and simple shapes. Furthermore, since the original motivation of this research was to determine whether or not demand-responsive transit can be used to improve transit satisfaction for passengers over traditional, fixed-route transportation while also minimizing costs to the operator, a series of simulations around MARTA stations is performed.

In these simulations, passenger costs and operator costs are assigned monetary values. According to the National Transit Database, the average per mile cost of operating fixed-route transit in the United States is $9.60 per mile [68] [69]. This value represents the total cost of operating fixed-route transit across all reporting agencies divided by the total number of vehicle miles traveled by those agencies. The cost encapsulates vehicle costs, maintenance costs, driver costs and benefits as well as any other fixed-route related costs reported by those agencies.

The same National Transit Database document indicates that the average per mile cost of operating dial-a-ride or paratransit services is $4.40 per mile [68] [69]. Just as with the fixed-route per mile costs, this value incorporates all vehicle purchase costs, maintenance costs, and staff costs for operating demand-responsive transportation. The majority of the decrease in per mile cost for demand-responsive versus fixed-route vehicles is due to the type of vehicles being used. While fixed-route vehicles use large buses, demand-responsive vehicles are typically minibuses or vans.

The finding that the per mile cost of demand-responsive transportation is cheaper than fixed-route transportation is at odds with the generally accepted notion that demand-responsive transportation is more expensive that fixed-route transportation [15]. Demand-responsive transportation is more expensive that fixed-route transportation in terms of passenger miles. Demand-responsive transportation costs $3.60 per passenger mile while fixed-route bus transportation only costs $0.90 per passenger mile [68] [69]. This is due to the fact that demand-responsive vehicles usually provide service in rural areas or to passengers with mobility challenges. For this reason, demand-responsive vehicles often operate with relatively few passengers and over a larger area than fixed-route transit. The demand-responsive service simulated below will not suffer from the problem of servicing a small number of passengers over a large area.

The passenger cost is a function of the time each passenger spends in transit, waiting for transit, or walking to a transit stop. Quantifying the value of time (VoT), in terms of monetary value, is subjective in nature. However, studies of passenger behavior provide some guidelines to determining passenger VoT. In a study published by the American Highway Users Alliance, time spent traveling is estimated to be 50% of a passenger's hourly wage [70]. While this is value a rough estimate, it will be used as an approximation of each passenger's value of time. A monetary representation of each passenger's VoT is needed in order to make objective comparisons with the monetary cost levied upon the operator. For the City of Atlanta, the average wage is $22.80 per hour [71]. This means that the estimate of each passenger's VoT is $11.40 per hour.

However, not all passenger values of time should be treated equally. The total passenger cost function includes weighted costs for time spent walking, time spent waiting, and time spent in transit. In simulations where these costs are all treated equally, the result of the optimization function was to force passengers to walk for extremely long periods of time. Some of these walking times approached one hour in duration. The reason for this was that the optimization algorithm could save operator costs by having many passengers walk large portions of their trip. In reality, this would lead to extreme customer dissatisfaction and disuse of the service. A more accurate representation would place a much larger penalty for having passengers walk long distances [72]. In a study by the Victoria Transport Policy Institute, passenger value of time spent walking was shown to vary from 1.5 to 4 times the value of time spent in the vehicle, depending on sidewalk conditions and other factors [73]. Therefore, in order to create a better representation of passenger behavior, the value of passenger walk time is set to three times the value of transit time in the following simulations.

The passenger and operator costs described above will be used to optimize the size and shape of a demand-responsive feeder zone at Midtown Station in Atlanta, GA. Table 6-3 lists the costs for convenience.

Table 6-3: Monetary costs for passengers and transit operators.

Fixed-Route Bus Costs	$9.60 per mile [68]
Demand-Responsive Vehicle Costs	$4.40 per mile [68]
Baseline Passenger Value of Time	$11.40 per hour [70] [71]

6.4.2 Feeder Zone Test Case Results

In this section, test cases are run to ascertain the feasibility of using isochrones to define demand-responsive feeder areas as well as to determine whether or not demand-responsive transportation can be used to improve performance in low-density urban areas.

Test Case 1: Comparing Isochrones and Ring-Radial Feeder Zones

In the first test case, the demand-responsive simulator will be used to determine if demand-responsive feeder zones defined by isochrones can provide better performance than feeder zones defined by radii. In this test case, a demand-responsive zone is proposed for Midtown Station in Atlanta, Georgia. The feeder zone will have two boundaries. The outer boundary represents the maximum distance that a demand-responsive vehicle will travel away from Midtown Station. The inner boundary represents the maximum distance that a passenger is expected to walk. Any passengers requesting trips within this boundary will be required to walk to the transit station.

The test case simulates passenger requests near the Midtown station on a weekday between 11:00 A.M. and 2:00 P.M. The costs of all passengers with origins or destinations within the feeder zone will be considered in this test case. The list of passengers is derived from the Atlanta Regional Commission On-board Passenger Survey [65]. Fixed-route transportation will be used for any portion of the trip that does not fall within the demand-responsive coverage area. The coverage area is intended to service the first or last mile of each passenger's trip.

Midtown Station on a weekday is one of the more heavily-trafficked stations in Atlanta. This time and location were chosen intentionally in order to push the limits of demand-responsive transportation and provider a greater separation between the performances of the isochrone-based setup and the radius-based setup. It is not expected that demand-responsive transportation will outperform fixed-route transportation in this case. This test case is used to show that using isochrones to represent the boundaries of the service areas can provide better results than using a straight-line distance such as a radius. To place further burden on the system, nearby transit stations are not considered as potential gateways in the test. Any passenger with an origin or destination within the coverage area will be sent to Midtown station, even if another station may be more efficient.

For each type of coverage area, a wide range of sizes was tested. As the size of the coverage areas increased, the number of passengers eligible for service increased. Figure 6-11 shows the net costs for implementing radius-based and isochrone-based coverage areas mapped against the number of passengers that they served. The total net cost is derived from Equation 6.12 and uses the costs defined in Table 6-3.

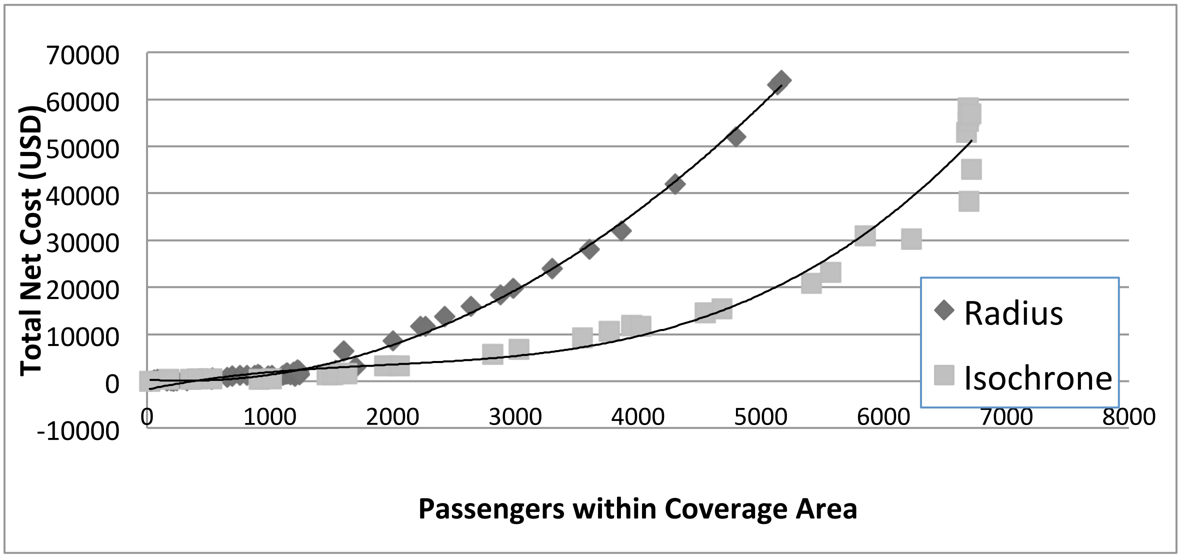

Figure 6-11: The total net cost of implementing increasingly large radius-based and isochrone-based coverage areas.

Figure 6-11 clearly shows that isochrone-based coverage areas are much more efficient in this scenario. As the coverage areas grow in size, and the number of passengers within the area increases, the cost of implementing radius-based coverage areas increases much faster than the cost of implementing isochrone-based coverage areas. This is the expected result. Using a radius to determine where the boundary will be only considers straight-line distance as the determining factor for whether a passenger is within the coverage area or not. A short straight line distance does not always correspond to a short driving time. The isochrone-based coverage areas will eliminate areas that are time-consuming to reach while including additional areas that may be outside the radius but less time-consuming to reach.

It should be noted that both isochrone and radius-based coverage areas fail to outperform fixed-route transit in this scenario. This is an expected result not only because Midtown is a very busy station on weekdays, but also because surrounding transit stations were not considered in these simulations. If the coverage areas were large enough to envelope surrounding transit stations, those stations were not considered as potential gateways. In a more practical implementation, they would be included in the optimization algorithm. These stations were ignored in this test case to reduce the number of optimization variables in order to focus solely on the effect of using isochrone and radius-based coverage areas to best serve a singular gateway.

Given this evidence that isochrone-based coverage areas show improvement over radius-based coverage areas, the next logical step is to determine if isochrone-based coverage areas can improve performance over fixed-route service.

Test Case 2: Optimizing Demand-Responsive Performance at Chamblee Station

In the second test case, the size and shape of a demand-responsive feeder system is optimized at Chamblee Station on a Sunday between 11:00 A.M. and 2:00 P.M. This time and location were chosen because the demand during this time is relatively low, making it a good candidate to operate demand-responsive transportation instead of fixed-route transportation. To further improve the performance of the demand-responsive system, adjacent transit stations are taken into account when assigning passengers to coverage areas. If the passenger is closer to any station besides Chamblee Station, then the passenger is not considered to be in the Chamblee coverage area. All passenger origins and destinations are restricted to be serviced by the transit station that is nearest them in terms of driving time. Assigning this limitation immediately creates a hard boundary for the Chamblee Station coverage area. Figure 6-12, shows this boundary. Any passenger that is not within the solid line is considered to be outside the Chamblee Subnet.

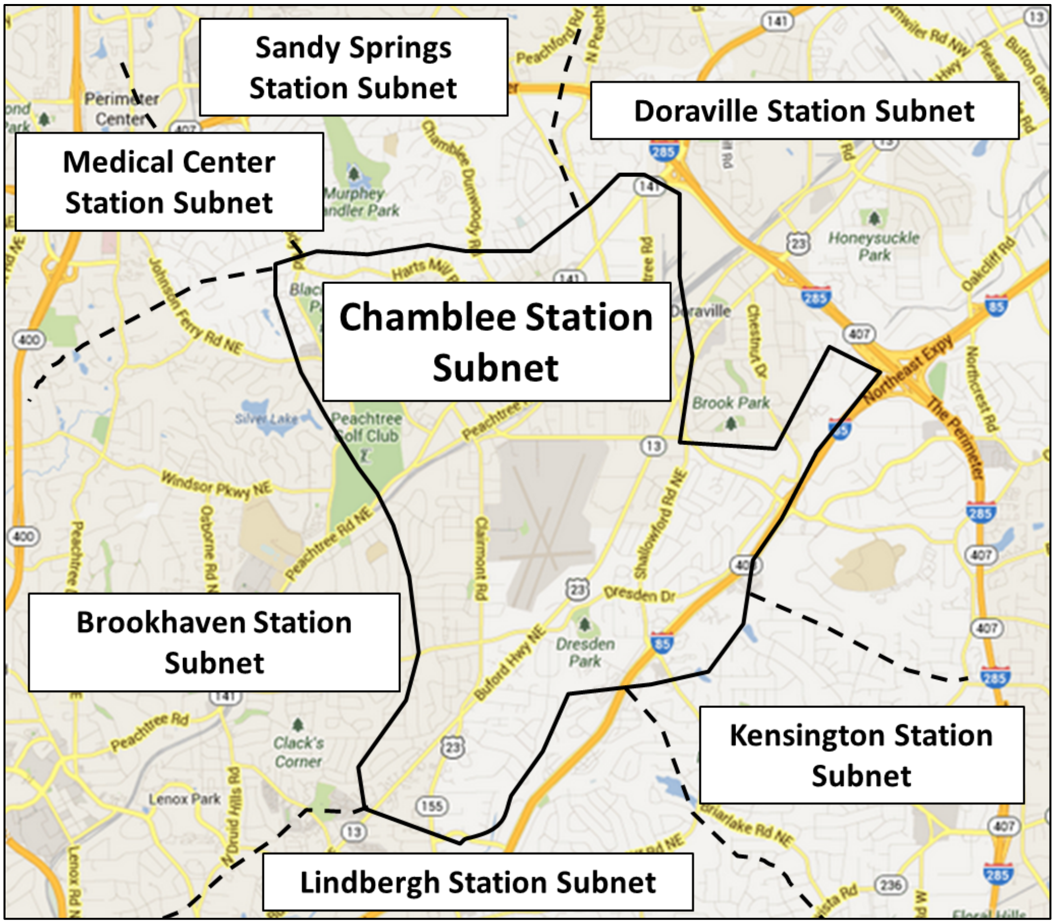

Figure 6-12: The Chamblee Station subnet lies within the solid black line. All points within this line are closer, in terms of driving time, to the Chamblee MARTA station than any other station.

Given this additional restriction, defining the optimal inner and outer isochrones can proceed as normal. Just as with Test Case 1, the particle swarm optimizer is used to test a range of possible isochrones. Using the comparison technique of Chapter 5, a specific walking isochrone size and driving isochrone size are tested. After each simulation, the total passenger and operator costs, derived from Table 6-2, are calculated. Based on the results, each particle is updated with new isochrones times and the simulations are repeated. Figure 6-13, shows the results of several iterations of particle simulations and updates. Each circle in this figure represents a simulation and comparison between a given demand-responsive setup and the existing fixed-route network. The size of the circle indicates the net cost of implementing a given demand-responsive simulation. Net cost refers to the cost of implementing a demand-responsive feeder zone minus the cost of using fixed-route transit. Positive costs indicate net increases in cost and negative costs indicate a net savings. Any setup with a negative cost indicates a setup in which demand-responsive transportation is more efficient than fixed-route transportation.

The location of the circle within the grid indicates the size of the driving and walking isochrones. The y-axis indicates the walking-isochrone size, and the x-axis indicates the driving isochrone size. The circle at position (0,0) represents a net cost of \$0. At (0,0), no demand-responsive system is present, and the entire system is operated as a fixed-route system. Any circle smaller than this circle indicates a setup where the demand-responsive system shows an improvement over the fixed-route system. The minimum net cost across all the simulations was found to be -\$1337 and was achieved with a walking isochrone size of 0 seconds and a driving isochrone size of 331 seconds.

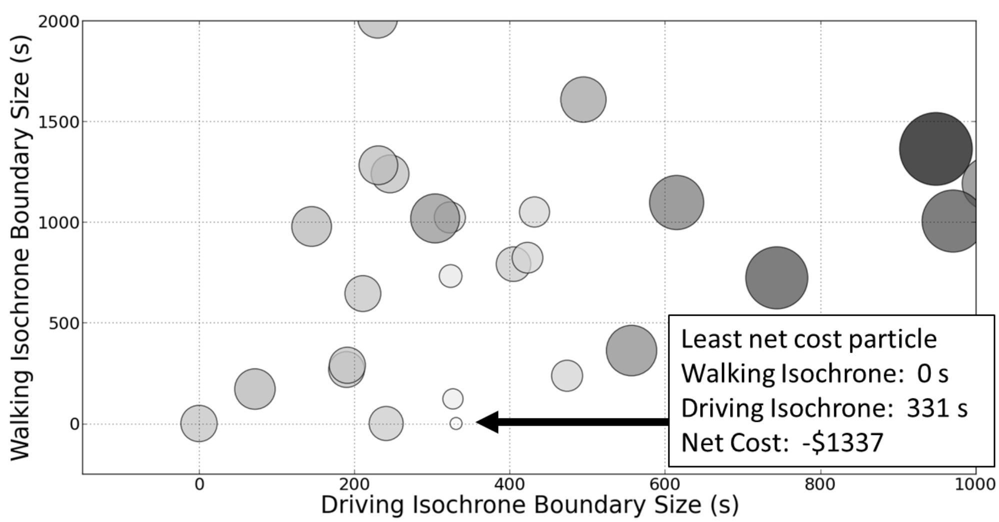

Figure 6-13: The relative costs of varying the isochrone times in the Chamblee Coverage area. The minimum cost particle was found when the driving isochrone is set to 331 seconds, and the walking isochrone is set to 0 seconds.

Figure 6-13 provides information about how sizing the isochrones affects performance. As the driving isochrone size increases, the operator costs also increase. This is indicated by the increasingly large circles on the right side of the graph. Similarly, as the walking isochrone grows larger, travelers are forced to walk longer distance, increasing the net passenger costs. The least costly particles, in the middle of the graph, indicate the best tradeoffs between passenger satisfaction and operator costs.

The net cost of -$1337 was achieved largely through decreasing passenger door-to-door travel time and decreasing passenger walk time. Many bus routes within the Chamblee coverage area operate at very high headways on Sunday. Some headways approach 1 hour in duration. The means that each time a traveler wants to depart, he or she must wait an average of 30 minutes before leaving. This overhead can have a dramatic impact on the total time between when a traveler wishes to depart and when the traveler actually arrives at the final destination. The average duration between a passenger's desired departure time and the actual arrival time for passengers whose trips begin or end within the Chamblee subnet is 2 hours and 0 minutes when fixed-route transit it used. In the optimal demand-responsive setup found by the particle swarm optimizer, the average duration is 1 hour and 20 minutes. Much of this time savings comes from eliminating the long headways between fixed-route buses. The tradeoff to the operator is an increase in vehicle miles traveled (VMT). Fleet VMT within the zone increased from 126.8 miles to 236.7 miles. However, the increase in VMT is offset by replacing large buses with smaller paratransit vehicles. Using the operating costs from Table 6-3, the increase in VMT actually comes at a small decrease in operating price. The net operating price was -$175.

Ultimately, the tradeoff between increasing passenger satisfaction versus operator costs or savings is a policy decision. Transit engineers will need to determine if increasing passenger satisfaction will lead to an increase in revenue through higher ridership. Alternatively, transit engineers can set limits on operator costs in order to see what the effect will be on passenger satisfaction. The demand-responsive simulator and particle swarm optimizer is intended as a tool to transportation professionals to test various pricing and operating strategies. There will always be a tradeoff between improving passenger satisfaction and limiting operator costs. The approach presented here helps to inform the decision on where to strike the balance between those two competing objectives.

6.5 Summary

The work presented in this chapter introduces a method of optimizing demand-responsive coverage areas in transit feeder systems. The particle swarm optimization algorithm demonstrated in this chapter relies upon the comparison techniques introduced in Chapter 5 and takes advantage of isochrones to determine optimal boundaries for feeder zone coverage areas.

The combination of isochrone-based boundaries, objective comparison of heterogeneous transportation systems, and particle swarm optimization allowed for an accurate simulation of an optimized demand-responsive feeder network at Chamblee MARTA Station. Using realistic values for vehicle costs, driver costs, and passenger values of time, it was demonstrated that a demand-responsive transportation system could provide better results at this transit station that the current fixed-route network.

More importantly, the methods introduced in this chapter can be easily applied to any transit network. Using widely available open-source data to represent transit schedules and road networks, the optimization method and simulator built for this research can be easily adapted to any situation. Transit engineers can use this approach to determine if a demand-responsive feeder network can improve customer satisfaction and save operator costs in their cities. If it is determined, though comparison, that a demand-responsive system can improve performance, the optimization techniques introduced in this research will assist the transit engineers and planners in implementing the most optimal system for their particular situation.

CONCLUSION

The original motivation for this research was to determine if demand-responsive transportation could be integrated with, or substituted for, fixed-route transit in low-density urban areas that lack the ridership demand to support frequent fixed-route transit. A review of the current state-of-the art in this area revealed that advancements had been made in comparing fixed-route transportation to demand-responsive transportation, and methods had been developed to optimize demand-responsive transportation coverage areas. However, much of this research focused on homogeneous areas where street networks were gridded, transit vehicles arrived at regular and unchanging intervals, and passengers were assumed to be evenly distributed both temporally and spatially. With these assumptions, analytical models representing passenger costs and operator costs could be derived. However, since the low-density urban areas that this research intended to study often lack the homogeneity required by the existing methods, a new approach is needed.

Instead of deriving analytical formulas to compare and optimize demand-responsive transportation systems, this research introduces a simulation-based approach to analyze demand-responsive transportation performance and optimize demand-responsive coverage areas in heterogeneous environments. The simulation-based comparison and optimization approaches take advantage of several recent contributions. Perhaps the most important contributions come from Quadrifoglio, Li, and others who developed the basic framework for comparing and optimizing demand-responsive transportation. The simulation-based approach used in this research is an adaptation of their analytical approaches in order to allow for comparisons in heterogeneous environments.

Another recent contribution allowing for this research is the proliferation of open-source projects that provide transit schedule information, optimal transit routing, and optimal street network routing. These projects, which include the OpenTripPlanner, Open Street Map, and the Open Source Routing Machine, provide the fundamental routing building blocks that make up the demand-responsive simulator built for this research.

This research extends the previous work in comparing and optimizing demand-responsive transportation systems to a wider range of environments. The major technology contribution of this research is the demand-responsive transportation simulator and the ability to compare and optimize transportation systems in any type of environment as long as the map, schedule, and passenger data for that environment is available. Given realistic passenger data, transit schedules in the GTFS format, and street map data, the demand-responsive simulator can simulate each and every passenger trip and vehicle itinerary during a given study period. The important contribution is that analytical formulas to represent passenger walk, wait, and transit times as well as operator costs no longer need to be derived for each and every comparison between transit systems. Furthermore, in areas where deriving the formulas is not practical or where accurate representations of costs cannot be found, the simulator provides an accurate and simple method of comparing transportation performance.

This research illustrated the comparison method in the city of Atlanta, GA. The comparison method looked at the possibility of implementing a radius-based demand-responsive feeder zone around every MARTA station in Atlanta. The results of the comparison showed that for most cases, fixed-route transit was superior to demand-responsive transit in Atlanta. However, further optimization of the demand-responsive layout could lead to improvements in demand-responsive performance.

A method of optimizing the size and shape of demand-responsive feeder areas is an additional major contribution of this work. The optimization scheme introduced in this research uses a particle swarm optimization approach and takes advantage of the demand-responsive transportation simulator to determine exactly what shape and how large demand-responsive feeder areas should be. The particle swarm optimizer would optimize the size of the outer and inner boundaries of the feeder zone and demonstrated through simulation that using isochrones to define these boundaries was superior than using a radius. By replacing the radius-based coverage areas at Chamblee station in Atlanta with isochrone-based coverage areas, and by using the particle swarm optimizer to continually test and improve the sizes of the boundaries, an optimal size and shape for the feeder zone was found. Through simulation it was conclusively demonstrated that demand-responsive transportation at Chamblee station on a weekend is more efficient, in terms of combined passenger and operator costs, than fixed-route at the same location. This conclusion begins to answer the original motivating question of this research as to whether demand-responsive transportation can be more efficient than fixed-route transportation in low-demand areas.

The creation of the demand-responsive transportation simulator and the adaptation of the comparison and optimization techniques allows transportation engineers to make informed decisions about where and how to implement a demand-responsive transportation scheme. However, further research is needed to fully implement the comparison and optimization methods in a practical setting. This future research should focus on handling the interaction between fixed-route and demand-responsive transportation. The simulations in this research assumed no interaction between fixed-route and demand-responsive vehicles at the boundaries of the demand-responsive feeder zones. In reality, the passengers outside these zones would likely need to access locations within the zone, and determining the best method for handing passenger hand offs between fixed-route and demand-responsive vehicles needs to be determined. Further improvements could consider the effects of other forms of demand-responsive transportation. This research focused on paratransit, however future work should consider the effect of ridesharing or deviated-route service in designing optimal demand-responsive coverage areas. Still further studies should be done to determine how to transition between fixed-route and demand-responsive transportation. What will the effect be on ridership? How will passengers without access to internet-capable devices make trip requests?

This research provides a strong first step towards demonstrating how demand-responsive transportation can be used to improve customer satisfaction and save operating costs in low-density urban areas. As more practical questions are answered, the work performed for this thesis will allow transportation engineers to accurately assess the viability of demand-responsive transportation in their system and assist in determining exactly how such a system should be implemented to meet global system objectives.

APPENDIX

NETWORK-INSPIRED TRANSPORTATION SIMULATOR USER'S GUIDE

This research required that a custom simulator be built in order to simulate passenger trips occurring in heterogeneous settings while minimizing assumption made about transit schedules, road layouts, and passenger behavior. The NITS Simulator can be found at, https://github.com/dedwards8/NITS.

The NITS Simulator is based on the Django framework and will require Django to be installed. This guide will not provide details on setting up Django or installing the required packages. Instructions for installing Django can be found at http://www.djangoproject.com, and the additional packages required by the NITS simulator are found in requirements.txt in the top level of the code repository.

The NITS simulator has a web-based graphical user interface. Once the Django environment is setup and the web server is activated, many of the main functions of the simulator are accessed via web pages. A list of commonly used URLs is defined below. In these URLs, <host address> refers to the web address of the root NITS Simulator Django project.

1. <host address>/*startanewsimulation*: Navigating to this address will start a new NITS simulation.

2. <host address>/*results*: Visit this page to view exact paths of vehicle and a summary of passenger and vehicle costs.
3. <host address>/*summary*: This page provides a summary of passenger and vehicle costs as well as map of passenger starting and ending locations.
4. <host address>/*subnets*: This page provides information on all available subnets and allows the user to create or delete subnets.
5. <host address>/*PSO*: This page manages the particle swarm optimization routine that finds the optimal isochrone sizes for each subnet.

In future versions of this project, these URLs will be combined into a more user-friendly central GUI.

Configuring the NITS Simulator

After installing the Django environment, downloading the NITS Simulator from GitHub, and importing the project into Django, the *settings.py* file should be customized to attain the desired operation. All changes required to setup a NITS demo for any region are made within the *settings.py*. In the first releasable version, this file will be accessed via GUI, however changes must now be made directly to the *NITS_CODE/settings.py* file.

The most important settings that need to be configures are:

1. SURVEY_PASSENGER_FILE: This value should point a comma separate file that contains trip information for one passenger on each line. All passengers that will request trips during the simulation are stored in this file. The format of each line is,

   [Passenger Id, Seconds, Start Lat, Start Lng, End Lat, End Lng]

Passenger Id is a unique passenger identifier. Seconds is the number of seconds into the simulation that the passenger will request at rip, e.g., if a passenger requests a trip at 10:10 AM for a simulation that starts at 10AM, the Seconds entry for that passenger will be 600. The starting and ending latitude and longitude of the passenger is stored in the final four values.

2. OTP_SERVER_URL: This value should be the URL of an OpenTripPlanner instance that provides optimized transit trips for the study area. Instructions for setting up an instance of OpenTripPlanner can be found at https://github.com/openplans/OpenTripPlanner/wiki/FiveMinutes.
3. OSRM_SERVER_URL: This value should be the URL of an Open Source Routing Machine server that provides optimal point-to-point vehicle trips for the study area. Instructions for setting up a custom instance of the Open Source Routing Machine can be found at http://project-osrm.org/.

These three settings must be set in order for the simulator to operate. Additional settings that may need to be set include the following.

4. PASSENGER_VOT: Passenger Value of Time. This value represents the passenger's value of time in terms of USD per hour. It is used in the route and itinerary optimization algorithms.
5. FRT_CPM: Fixed-route transportation cost per mile. This is the average cost of operating a fixed-route vehicle based on distance. It measured in USD/meter. It is used in the route and itinerary optimization algorithms.

6. DRT_CPM: Demand-responsive transportation cost per mile. This is the average cost of operating a demand-responsive vehicle based on distance. It is measured in USD/meter. It is used in the route and itinerary optimization algorithms.
7. CHECK_DRIVING_TIME: If this is set to true, the outer boundary of each subnet will be defined by a maximum driving time.
8. CHECK_WALKING_TIME: If this is set to true, the inner boundary of each subnet will be defined by a maximum walking time.
9. CHECK_RADIUS: If this is set to true, the boundaries of each subnet will be defined by straight line distances and not isochrones.
10. SIMULATION_START_TIME: This is the time of day that the simulation will start. It is measured in seconds from midnight.
11. SIMULATION_START_DAY, SIMULATION_START_MONTH, SIMULATION_START_YEAR: These values define the exact day that is being simulated. This information is used by the transit operator to provide transit itineraries specific to each day.
12. USE_PSO: If this is set to true, the simulation will not stop after one iteration. After each simulation period is finished, another simulation will be started based with specifications set by the Particle Swarm Optimizer.

Creating the Subnets

The final piece of required information needed by the NITS simulator is the subnet information. A subnet is defined by a gateway, and inner boundary, and an outer boundary. Subnets are defined by navigating any browser to the <host address>*/subnets/* URL. Here a list of all the subnets is listed. Subnets can be deleted or added to this list.

To add a subnet, define a latitude for the gateway, a longitude for the gateway, a maximum driving distance for the isochrone, and a maximum walking distance for the isochrone. When the simulation begins any passenger that falls within one of these subnets will inserted into a demand-responsive vehicle and routed to the gateway of that subnet. If no subnets are defined, or no passengers fall within a subnet, then all trips will be fixed-route transit trips.

Running the Simulation

Once all the *settings.py* settings are defined and the subnets are created, a simulation can be started by navigating to <host address>*/startanewsimulation/*. This page manages all the high level actions of the simulator. After navigating to the page, all previous simulation information is cleared and a new simulation is started.

Managing the Particle Swarm Optimizer

The particle swarm optimizer is managed by navigating to <host address>/*PSO*/. From this screen the current status of each particle is displayed. The user can kick off a new optimization routing here by re-initializing all particles. The user can individually change particles to test specific isochrone values. From this screen, the user also defines which subnet is being tested. The particle swarm optimizer only optimizes one subnet at a time. So while many subnets may exist in the system, the values of each particle only dictate the behavior of a single subnet. The value *x1* is the size of the outer boundary and the value *x2* is the size of the inner boundary. If the boundaries are defined using isochrones, then these values represent travel time from the gateway, measured in seconds. If the boundaries are defined using radii, then these values represent haversine distance from the gateway and they are measured in meters.

It is important to note that in order to the particle swarm optimizer to calculate the total cost of operating fixed-route transit within a given subnet, general transit feed specification data must be uploaded into the simulator. The GTFS data should be the same data used by the OpenTripPlanner. GTFS data for many agencies can be found at www.gtfs-datta-exchange.com . To insert the GTFS data into the NITS database, unzip the GTFS file into the */hermes/bin* directory and execute the */hermes/gtfs_insert.py* script.

Customizing the Software

If it is desired to alter the inner workings of the simulator, such as updating the particle swarm optimization algorithm or the dial-a-ride route selection algorithm, there are five main files that control the operation of the NITS simulator. The five files are *master.py*, *passenger_manager.py*, *views.py*, *results.py*, and *particle_swarm_manager.py*. A basic flow chart of how these files interact is shown in Figure A-1.

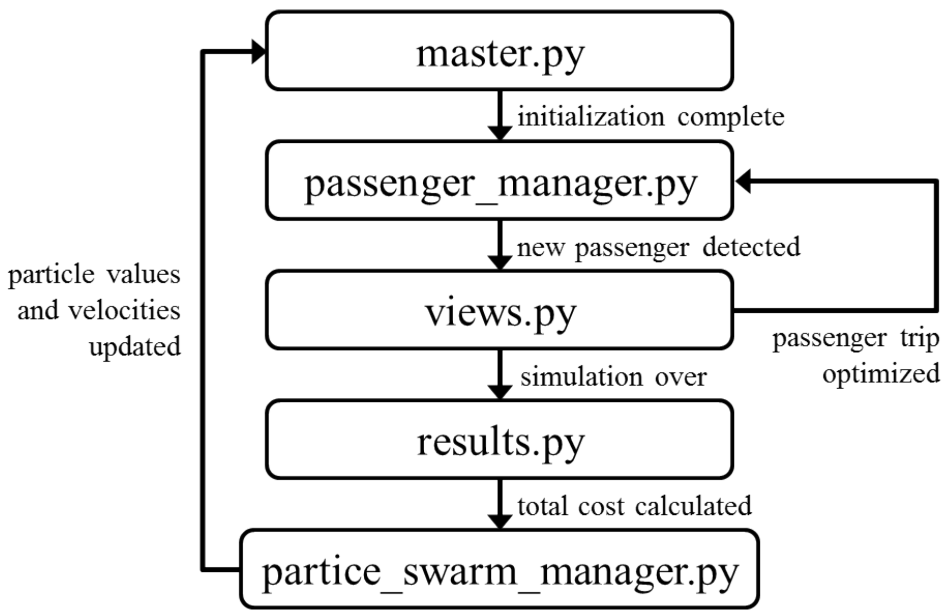

Figure A-1: Flow of data within the NITS Software Simulator

The basic function of each of these files is as follows.

master.py: This file initializes each simulation. All the vehicle data is reset and all passenger trips are deleted. If a subnet is being optimized by the particle swarm optimizer, the boundaries for that subnet are updated here.

passenger_manager.py: At each second, *passenger_manager.py* checks for passengers in the passenger survey database. If a new passenger is created that passenger is passed to *views.py* where trips are created and optimized. *passenger_manager.py* also checks for previously created trips that are ready to start. For instance, if a passenger requires two demand-responsive trips to reach his or her final destination, *passenger_manager.py* will determine when the passenger is available to start the second trip.

views.py: In this file, trips are created and optimized. When a new trip is passed to *views.py*, the proper subnet for that trip is identified and the optimal vehicle and route is determined. *views.py* also manages transit and walking trips.

results.py: This is where all the passenger and operator costs are tallied. For each passenger the walking, waiting, and riding times are summed and multiplied by the applicable cost values. Demand-responsive vehicle routes are summed and multiplied by the per-mile cost of operating demand-responsive vehicles. Fixed-route costs are calculated by summing all the trip shape files that fall within the subnet coverage area during the simulation time. This value is multiplied by the per-mile cost of operating fixed-route transit.

particle_swarm_manager.py: Once the results are tallied in *results.py*, the particle swarm optimizer updates the cost of the particle being tested, adjusts velocities of that particle, and loads the next particle for testing. The *particle_swarm_manager.py* then calls the *master.py* to run the next simulation.

REFERENCES

[1] B. S. Pushkarev and J. M. Zupan, Public Transportation and Land Use Policy, Indiana University Press, 1977.

[2] R. Cervero and K. Kockelman, "Travel Demand and the 3Ds: Density, Diversity, and Design," *Transportation Research Part D,* vol. 2, no. 3, pp. 199-219, 1997.

[3] United States Census, "Summary of Population and Housing Characteristics," 2010. [Online]. Available: http://www.census.gov/prod/www/decennial.html. July 12, 2013.

[4] A. Murray and X. Wu, "Accessibility Tradeoffs in Public Transit Planning," *Journal of Geographical Systems,* vol. 5, no. 1, pp. 93-107, 2003.

[5] TRB Transit Cooperative Research Program, Transit Capacity and Quality of Service Manual—2nd Edition, Washington, D.C.: Transportation Research Board, 2003.

[6] Metro Atlanta Rapid Transit Authorty, "MARTA Mobility Policy," 1 August 2012. [Online]. Available: http://www.itsmarta.com/accessibility-Mobility.aspx.

[7] T. Crainic, F. Malucelli, M. Nonato and F. Guertin, "Meta-Heuristics for a Class of Demand-Responsive Transit Systems," *INFORMS Journal on Computing,* vol. 17, no. 1, pp. 10-24, 2005.

[8] Georgia Institute of Technology, "Georgia Tech Paratransit Service," 1 June 2013. [Online]. Available: http://pts.gatech.edu/ride/Pages/Stingerette.aspx.

[9] Duke University Parking and Transportation, "Duke Vans," 1 June 2013. [Online]. Available: http://parking.duke.edu/buses_vans/duke_vans.php.

[10] Case Western Reserve University , "Safe Ride," CWRU Police Department, 1 June 2013. [Online]. Available: http://police.case.edu/saferide.html.

[11] Denver Regional Transportation District, "Denver RTD Call-n-Ride," Denver RTD, 1 June 2013. [Online]. Available: http://www.rtd-denver.com/callNRide.shtml.

[12] KutsuPlus, 20 October 2013 [Online]Available: https://kutsuplus.fi/.

[13] "Flex 401 Route," 2011.

[14] "MyBus Rural," Available: http://www.southlanarkshire.gov.uk, July 1, 2013.

[15] J.-F. Cordeau and G. Laporte, "The dial-a-ride problem: models and algorithms," *Annals of Operations Research,* vol. 153, no. 1, pp. 29-46, 2007.

[16] L. Fu and G. Ishkhanov, "Fleet Size and Mix Optimization for Paratransit Services," *Transportation Research Record: Journal of the Transportation Research Board,* vol. 1884, pp. 39-46, 2004.

[17] Z. Zhang and W. Cai, "A dynamic shortest path algorithm based on real-time traffic information in the urban public transit network," in *Service Operations and Logistics, and Informatics*, 2008.

[18] J. Jaw, A. Odoni, H. Psaraftis and N. Wilson, "A Heuristic Algorithm for the Multi-Vehicle Advance Request Dial-a-Ride Problem with Time Windows," *Transportation Research Part B,* vol. 20, pp. 243-257, 1986.

[19] T. Sexton and L. Boudin, "Optimizing single vehicle many-to-many operations with desired delivery times: Routing," *Transportation Science,* vol. 19, pp. 411-435, 1985.

[20] K. Uchimura, H. Takahashi and T. Saitoh, "Demand Responsive Services in Hierarchical Public Transportation System," in *IEEE Transactions on Vehicular Technology*, July, 2002.

[21] R. Jorgensen, J. Larsen and K. Bergvinsdottir, "Solving the Dial-a-Ride problem using genetic algorithms," *Journal of the Operational Research Society,* vol. 58, pp. 1321-1331, 2007.

[22] M. Aldaihani and M. Dessouky, "Hybrid scheduling methods for paratransit operations," *Computers and Industrial Engineering,* vol. 45, no. 1, pp. 75-96, 2003.

[23] O. Madsen, H. Ravn and J. Rygaard, "A heuristic algorithm for a dial-a-ride problem with time windows, multiple capacities, and multiple objectives," *Annals of Operations Research,* vol. 60, pp. 193-208, 1995.

[24] J. qing Wang, X. nian Tong and Z. mao Li, "An Improved Evolutionary Algorithm for Dynamic Vehicle Routing Problem with Time Windows," in *Lecture Notes in Computer Science*, vol. 4490, Y. Shi, G. D. van Albada, J. Dongarra and P. Sloot, Eds., Springer-Verlag, 2007, pp. 1147-1153.

[25] T. Miyamoto, K. Nakatyou and S. Kumagai, "Route Planning Method for a Dial-a-Ride Problem," in *IEEE International Conference on Systems, Man and Cybernetics*, 2003.

[26] N. Wilson, R. Weissberg and J. Hauser, "Advanced dial-a-ride algorithms research report: Final Report," 1976.

[27] C.-F. Liaw, C. White and J. Bander, "A Decision Support System for the Bimodal Dial-A-Ride Problem," in *IEEE Transactions on Systems, Man, and Cybernetics*, 1996.

[28] C. Hall, H. Andersson, J. Lundgren and P. Varbrand, "The Integrated Dial-a-Ride Problem," *Public Transport,* vol. 1, no. 1, pp. 39-54, 2009.

[29] K.-T. Lee, P.-J. Wu and S.-H. Wang, "The Planning and Design of Taxipooling on Feeder System," in *Proceedings of the 2004 IEEE International Conference on Networking, Sensing and Control*, 2004.

[30] M. Horn, "Procedures for planning multi-leg journeys with fixed-route and demand-responsive passenger transport services," *Transportation Research Part C: Emerging Technologies,* vol. 12, no. 1, pp. 33-55, 2004.

[31] M. Diana, L. Quadrifoglio and C. Pronello, "A Methodology for Comparing Distances Traveled by Performance-Equivalent Fixed-Route and Demand Responsive Transit Services," *Transportation Planning and Technology,* vol. 32, no. 4, pp. 377-399, 2009.

[32] R. Thompson, K. Sakulchariyalert, J. Haasz, S. Winter and P. Mendis, "Determining the Viability of a Demand Responsive Transport System".

[33] X. Li and L. Quadrifoglio, "Optimal Zone Design for Feeder Transit Services," *Transportation Research Record: Journal of the Transportation Research Board,* vol. 2111, pp. 100-108, 2009.

[34] L. Q. Xiugang Li, "Feeder Transit Services: Choosing between fixed and demand responsive policy," *Transportation Research Part C,* vol. 18, pp. 770-780, 2010.

[35] L. Quadrifoglio, "A Hybrid Fixed and Flexible Transportation Service: Description, Viabililty, Formulation, Optimization, and Heuristic," 2005.

[36] L. Quadrifoglio and X. Li, "A Methodology to Derive the Critical Demand Density for Designing and Operating Feeder Transit Services," *Transportation Research Part B,* vol. 43, pp. 922-935, 2009.

[37] X. Li and L. Quadrifoglio, "2-Vehicle Zone Optimal Design for Feeder Transit Services," *Public Transport,* vol. 3, pp. 89-104, 2011.

[38] OpenStreetMap, "OpenStreetMap: The Free Wiki World Map," 1 August 2012. [Online]. Available: http://www.openstreetmap.org.

[39] P. Eklund, S. Kirkby and S. Pollitt, "A dynamic multi-source Dijkstra's algorithm for vehicle routing," in *Australian and New Zealand Conference on Intelligent Information Systems*, 1996.

[40] S. Kim, M. Lewis and C. White, "Optimal Vehicle Routing with Real-Time Traffic Information," *IEEE Transactions on Intelligent Transportation Systems,* vol. 6, no. 2, pp. 178-188, 2002.

[41] S. Kim, M. Lewis and C. White, "State Space Reduction for Nonstationary Stochastic Shortest Path Problems with Real-Time Traffic Information," *IEEE Transactions on Intelligent Transportation Systems,* vol. 6, no. 3, pp. 273-284, 2005.

[42] E. Miller-Hooks and H. Mahmassani, "Least Expected Time Paths in Stochastic, Time-Varying Transportation Networks," *Transportation Science,* vol. 34, no. 2, pp. 198-215, 2000.

[43] P. Sanders and D. Schultes, "Engineering Fast Route Planning Algorithm," in *WEA'07 Proceedings of the 6th international conference on Experimental algorithms*, 2007.

[44] D. Schultes, *Route Planning in Road Networks,* Chapter 4, "Highway-Node Routing", 2008.

[45] D. Schultes, *Route Planning in Road Networks*, Chapter 6, "Transit-Node Routing", 2008.

[46] F. Y. zhen, W. Q. chun, L. D. min and J. F. chao, "An Improved Dijkstra Algorithm Used on Vehicle Optimization Route Planning," in *2nd International Conference on Computer Engineering and Technology*, 2010.

[47] Project Open Source Routing Machine, "Project OSRM," 1 August 2012. [Online]. Available: http://project-osrm.org/.

[48] Google Developers, "What is GTFS?," Google, 1 August 2012. [Online]. Available: https://developers.google.com/transit/gtfs/.

[49] J. Czebotar, "GTFS Data Exchange," 1 August 2012. [Online]. Available: http://www.gtfs-data-exchange.com/.

[50] R. Huang, "A Schedule-Based Pathfinding Algorithm for Transit Networks Using Pattern First Search," *Geoinformatica,* vol. 11, pp. 269-285, 2007.

[51] H. Kao, "Schedule-Based Path Planning: An Optimal-Time Algorithm," 2008.

[52] G. R. Mauri and L. A. N. Lorena, "Customer's Satisfaction in a Dial-A-Ride Problem," *IEEE Intelligent Transportation Systems Magazine,* vol. 1, no. 3, pp. 6-14, Fall 2009.

[53] J.-F. Cordeau and G. Laporte, "A tabu search heuristic for the static multi-vehicle dial-a-ride problem," *Transportation Research Part B,* vol. 37, pp. 579-594, 2003.

[54] J.-F. Cordeau, "A Branch-and-Cut Algorith for the Dial-A-Ride Problem," *Operations Research,* vol. 54, pp. 573-586, 2006.

[55] D. Edwards, A. Elangovan and K. Watkins, "Reaching Low-Density Urban Areas with the Network-Inspired Transportation System," *IEEE Intelligent Transportation Systems Conference,* 2012.

[56] S. Chien and P. Schonfield, "Optimization of Grid Transit System in Heterogeneous Urban Environment," *Journal of Transportation Engineering,* vol. 123, pp. 28-35, 1997.

[57] S. Chang and P. Schonfeld, "Optimization Models for Comparing Conventional and Subscription Bus Feeder Services," *Transportation Science ,* vol. 25, pp. 281-198, 1991.

[58] L. Q. M. D. Majid Aldaihani, "Network design for a grid hybrid transit service," *Transportation Research Part A,* vol. 38, pp. 511-530, 2004.

[59] B. Taylor, "Thinking Outside the Bus: Understanding User Perceptions of Waiting and Transferring in Order to Increase Transit Use," California PATH Research Report, Berkely, 2009.

[60] X. Li, "Optimal Design of Demand-Responsive Feeder Transit Services," 2009.

[61] L. Quadrifoglio, M. Dessouky and F. Ordonez, "A Simulation Study of Demand-Responsive Transit System Design," *Transportation Research Part A,* vol. 42, pp. 718-737, 2008.

[62] L. Pages, R. Jayakrishnan and C. E. Cortes, "Real-Time Mass Passenger Transport Network Optimization Problems," *Transportation Research Record: Journal of the Transportation Research Board,* vol. 1964, pp. 229-237, 2006.

[63] Open Plans, "OpenTripPlanner," 1 August 2012. [Online]. Available: http://www.opentripplanner.com.

[64] D. Luxen, "Real-Time Routing with OpenStreetMap Data," *Proceedigns of the 19th ACM SIGSPATIAL International Conference on Advacnes in Geographic Information Systems,* vol. 11, pp. 513-516, 2011.

[65] Atlanta Regional Commission, "Atlanta Regional Commission On-Board Passenger Survey Final Report," June 2010. [Online]. Available: http://www.atlantaregional.com/File%20Library/Transportation/Travel%20Demand%20Model/tp_arcregionalonboardsurveyfinalreport_063010.pdf.

[66] J. Walker, Human Transit: How Clearer Thinking about Public Transit Can Enrich Our Communities and Our Lives, Washington, D.C.: Island Press, 2011.

[67] J. Kennedy and R. Eberhart, "Particle Swarm Optimization," *IEEE Internation Conference on Neural Networks, 1995 Proceedings,* vol. 4, pp. 1942-1948, 1995.

[68] Federal Transit Administration: US DOT, "National Transit Summaries and Trends," National Transit Database, Washington, D.C., 2011.

[69] Federal Transit Administration: US DOT, "Transit Profile: Top 50 Agencies for the 2011 Report Year," National Transit Database, Washington, D.C., 2012.

[70] T. Hogarty, Saving Time, Saving Money: The Economics of Unclogging America's Worst Bottlenecks, Washington, D.C.: American Highway Users Alliance, 2000.

[71] Bureau of Labor Statistics: U.S. Department of Labor, "Occupational Employment and Wages in Atlanta-Sandy Springs-Marietta, May 2012," U.S. Department of Labor, Southeast Information Office, Atlanta, GA, 2012.

[72] B. T. a. M. M. Hiroyuki Iseki, "The Effects of Out-of-Vehicle Time on Travel Behavior: Implications for Transit Transfers," California PATH Program, Institute of Tranportation Studies, University of California, Berkeley , 2006.

[73] T. Litman, "Valuing Transit Service Quality Improvements," *Journal of Public Transportation,* vol. 11, no. 2, pp. 43-63, 2008.